Little Mathematics Library

∞

A. S. SOLODOVNIKOV

SYSTEMS OF LINEAR INEQUALITIES

Mir Publishers . Moscow

ИЗДАТЕЛЬСТВО · МИР ·

ПОПУЛЯРНЫЕ ЛЕКЦИИ ПО МАТЕМАТИКЕ

А. С. Солодовников

СИСТЕМЫ ЛИНЕЙНЫХ НЕРАВЕНСТВ

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

LITTLE MATHEMATICS LIBRARY

A. S. Solodovnikov

SYSTEMS OF LINEAR INEQUALITIES

Translated from the Russian
by
Vladimir Shokurov

MIR PUBLISHERS
MOSCOW

First published 1979

Revised from the 1977 Russian edition

519.92
SHO

25.8.94
8714

На английском языке

© Издательство «Наука», 1977

© English translation, Mir Publishers, 1979

CONTENTS

Preface 7

1. Some Facts from Analytic Geometry 8

2. Visualization of Systems of Linear Inequalities in Two or Three Unknowns 17

3. The Convex Hull of a System of Points 22

4. A Convex Polyhedral Cone 25

5. The Feasible Region of a System of Linear Inequalities in Two Unknowns 31

6. The Feasible Region of a System in Three Unknowns 44

7. Systems of Linear Inequalities in Any Number of Unknowns 52

8. The Solution of a System of Linear Inequalities by Successive Reduction of the Number of Unknowns 57

9. Incompatible Systems 64

10. A Homogeneous System of Linear Inequalities. The Fundamental Set of Solutions 69

11. The Solution of a Nonhomogeneous System of Inequalities 81

12. A Linear Programming Problem 84

13. The Simplex Method 91

14. The Duality Theorem in Linear Programming 101

15. Transportation Problem 107.

Preface

First-degree or, to use the generally accepted term, *linear inequalities* are inequalities of the form

$$ax + by + c \geqslant 0$$

(for simplicity we have written an inequality in two unknowns x and y). The theory of systems of linear inequalities is a small but most fascinating branch of mathematics. Interest in it is to a considerable extent due to the beauty of geometrical content, for in geometrical terms giving a system of linear inequalities in two or three unknowns means giving a convex polygonal region in the plane or a convex polyhedral solid in space, respectively. For example, the study of convex polyhedra, a part of geometry as old as the hills, turns thereby into one of the chapters of the theory of systems of linear inequalities. This theory has also some branches which are near the algebraist's heart; for example, they include a remarkable analogy between the properties of linear inequalities and those of systems of *linear equations* (everything connected with linear equations has been studied for a long time and in much detail).

Until recently one might think that linear inequalities would forever remain an object of purely mathematical work. The situation has changed radically since the mid 40s of this century when there arose a new area of applied mathematics—*linear programming*—with important applications in the economy and engineering. Linear programming is in the end nothing but a part (though a very important one) of the theory of systems of linear inequalities.

It is exactly the aim of this small book to acquaint the reader with the various aspects of the theory of systems of linear inequalities, viz. with the geometrical aspect of the matter and some of the methods for solving systems connected with that aspect, with certain purely algebraic properties of the systems, and with questions of linear programming. Reading the book will not require any knowledge beyond the school course in mathematics.

A few words are in order about the history of the questions to be elucidated in this book.

Although by its subject-matter the theory of linear inequalities should, one would think, belong to the most basic and elementary parts of mathematics, until recently it was studied relatively little. From the last years of the last century works began occasionally to appear which elucidated some properties of systems of linear inequalities. In this connection one can mention the names of such mathematicians as H. Minkowski (one of the greatest geometers of the end of the last and the beginning of this century especially well known for his works on convex sets and as the creator of "Minkowskian geometry"), G.F. Voronoi (one of the fathers of the "Petersburg school of number theory"), A. Haar (a Hungarian mathematician who won recognition for his works on "group integration"), H. Weyl (one of the most outstanding mathematicians of the first half of this century; one can read about his life and work in the pamphlet "Herman Weyl" by I. M. Yaglom, Moscow, "Znanie", 1967). Some of the results obtained by them are to some extent or other reflected in the present book (though without mentioning the authors' names).

It was not until the 1940s or 1950s, when the rapid growth of applied disciplines (linear, convex and other modifications of "mathematical programming", the so-called "theory of games", etc.) made an advanced and systematic study of linear inequalities a necessity, that a really intensive development of the theory of systems of linear inequalities began. At present a complete list of books and papers on inequalities would probably contain hundreds of titles.

1. Some Facts from Analytic Geometry

1°. *Operations on points.* Consider a plane with a rectangular coordinate system. The fact that a point M has coordinates x and y in this system is written down as follows:

$$M = (x, y) \text{ or simply } M(x, y)$$

The presence of a coordinate system allows one to perform some operations on the points of the plane, namely *the operation of addition of points and the operation of multiplication of a point by a number.*

The addition of points is defined in the following way: if $M_1 = (x_1, y_1)$ and $M_2 = (x_2, y_2)$, then

$$M_1 + M_2 = (x_1 + x_2, \ y_1 + y_2)$$

Thus the addition of points is reduced to the addition of their similar coordinates.

The visualization of this operation is very simple (Fig. 1); the point $M_1 + M_2$ is the fourth vertex of the parallelogram constructed on the segments OM_1 and OM_2 as its sides (O is the origin of coordinates). M_1, O, M_2 are the three remaining vertices of the parallelogram.

The same can be said in another way: the point $M_1 + M_2$ is obtained by translating the point M_2 in the direction of the segment OM_1 over a distance equal to the length of the segment.

The multiplication of the point $M(x, y)$ by an arbitrary number k is carried out according to the following rule:

$$kM = (kx,\ ky)$$

The visualization of this operation is still simpler than that of the addition; for $k > 0$ the point $M' = kM$ lies on the ray OM, with $OM' = k \times OM$; for $k < 0$ the point M' lies on the extension of the ray OM beyond the point O, with $OM' = |k| \times OM$ (Fig. 2).

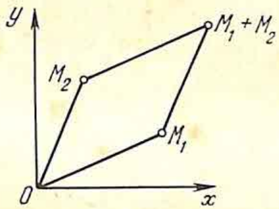


Fig. 1

The derivation of the above visualization of both operations will provide a good exercise for the reader*.

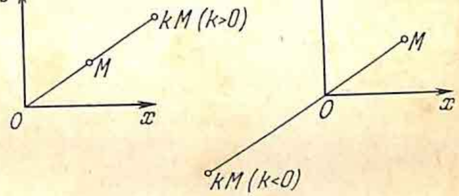


Fig. 2

The operations we have introduced are very convenient to use in interpreting geometric facts in terms of algebra. We cite some examples to show this.

---

* Unless the reader is familiar with the fundamentals of vector theory. In vector terms our operations are known to mean the following: the point $M_1 + M_2$ is the end of the vector $\vec{OM}_1 + \vec{OM}_2$ and the point kM is the end of the vector $k \times \vec{OM}$ (on condition that the point O is the beginning of this vector).

(1) *The segment M_1M_2 consists of all points of the form*

$$s_1M_1 + s_2M_2$$

where s_1, s_2 are any two nonnegative numbers the sum of which equals 1.

Here a purely geometric fact, the belonging of a point to the segment M_1M_2, is written in the form of the algebraic relation $M = s_1M_1 + s_2M_2$ with the above constraints on s_1, s_2.

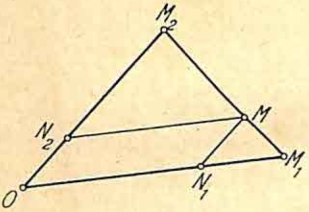


Fig. 3

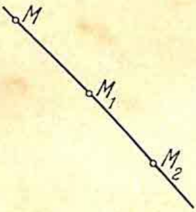


Fig. 4

To prove the above, consider an arbitrary point M on the segment M_1M_2. Drawing through M straight lines parallel to OM_2 and OM_1 we obtain the point N_1 on the segment OM_1 and the point N_2 on the segment OM_2 (Fig. 3). Let

$$s_1 = \frac{M_2M}{M_2M_1}, \quad s_2 = \frac{M_1M}{M_1M_2}$$

the numbers s_1 and s_2 being nonnegative and their sum equalling 1. From the similarity of the corresponding triangles we find

$$\frac{ON_1}{OM_1} = \frac{M_2M}{M_2M_1} = s_1, \quad \frac{ON_2}{OM_2} = \frac{M_1M}{M_1M_2} = s_2$$

which yields $N_1 = s_1M_1$, $N_2 = s_2M_2$. But $M = N_1 + N_2$, hence $M = s_1M_1 + s_2M_2$. We, finally, remark that when the point M runs along the segment M_1M_2 in the direction from M_1 toward M_2, the number s_2 runs through all the values from 0 to 1. Thus proposition (1) is proved.

(2) *Any point M of the straight line M_1M_2 can be represented as*

$$tM_1 + (1 - t)M_2$$

where t is a number.

In fact, if the point M lies on the segment M_1M_2, then our statement follows from that proved above. Let M lie outside of the segment M_1M_2. Then either the point M_1 lies on the segment MM_2 (as in Fig. 4) or M_2 lies on the segment MM_1. Suppose, for example, that the former is the case. Then, from what has been proved,

$$M_1 = sM + (1 - s)M_2 \quad (0 < s < 1)$$

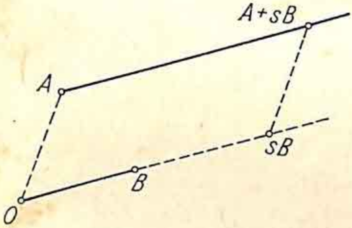


Fig. 5

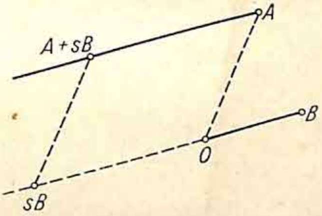


Fig. 6

Hence

$$M = \frac{1}{s}M_1 - \frac{1-s}{s}M_2 = tM_1 + (1-t)M_2$$

where $t = 1/s$. Let the case where M_2 lies on the segment MM_1 be considered by the reader.

(3) *When a parameter s increases from 0 to ∞, the point sB runs along the ray OB* and the point $A + sB$ is the ray emerging from A in the direction of OB. When s decreases from 0 to $-\infty$, the points sB and $A + sB$ run along the rays that are supplementary to those indicated above.* To establish this, it is sufficient to look at Figs. 5 and 6.

It follows from proposition (3) that, as s changes from $-\infty$ to $+\infty$, the point $A + sB$ runs along the straight line passing through A and parallel to OB.

The operations of addition and multiplication by a number can, of course, be performed on points in space as well. In that case, by definition,

---

* The point B is supposed to be different from the origin of coordinates O.

$$M_1 + M_2 = (x_1 + x_2,\ y_1 + y_2,\ z_1 + z_2)$$
$$kM = (kx,\ ky,\ kz)$$

All the propositions proved above will obviously be true for space as well.

We conclude this section by adopting a convention which will later help us formulate many facts more clearly and laconically. Namely, if $\mathscr{K}$ and $\mathscr{L}$ are some two sets of points (in the plane or in space), then we shall agree to understand by their "sum" $\mathscr{K} + \mathscr{L}$ a set of all points of the form $K + L$ where K is an arbitrary point in $\mathscr{K}$ and L an arbitrary point in $\mathscr{L}$.

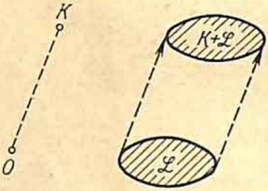


Fig. 7

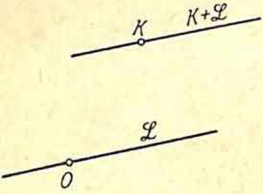


Fig. 8

Special notation has been employed in mathematics for a long time to denote the belonging of a point to a given set; namely, in order to indicate that a point M belongs to a set $\mathscr{M}$ one writes $M \in \mathscr{M}$ (the symbol $\in$ standing for the word "belongs"). So $\mathscr{K} + \mathscr{L}$ is a set of all points of the form $K + L$ where $K \in \mathscr{K}$ and $L \in \mathscr{L}$.

From the visualization of the addition of points a simple rule for the addition of the point sets $\mathscr{K}$ and $\mathscr{L}$ can be given. This rule is as follows. For each point $K \in \mathscr{K}$ a set must be constructed which is a result of translating $\mathscr{L}$ along the segment OK over a distance equal to the length of the segment and then all sets obtained in this way must be united into one. It is the latter that will be $\mathscr{K} + \mathscr{L}$.

We shall cite some examples.

1. Let a set $\mathscr{K}$ consist of a single point K whereas $\mathscr{L}$ is any set of points. The set $K + \mathscr{L}$ is a result of translating the set $\mathscr{L}$ along the segment OK over a distance equal to its length (Fig. 7). In particular, if $\mathscr{L}$ is a straight line, then $K + \mathscr{L}$ is a straight line parallel to $\mathscr{L}$. If at the same time the line $\mathscr{L}$ passes through the origin, then $K + \mathscr{L}$ is a straight line parallel to $\mathscr{L}$ and passing through the point K (Fig. 8).

2. $\mathcal{K}$ and $\mathcal{L}$ are segments (in the plane or in space) not parallel to each other (Fig. 9). Then the set $\mathcal{K}+\mathcal{L}$ is a parallelogram with sides equal and parallel to $\mathcal{K}$ and $\mathcal{L}$ (respectively). What will result if the segments $\mathcal{K}$ and $\mathcal{L}$ are parallel?

3. $\mathcal{K}$ is a plane and $\mathcal{L}$ is a segment not parallel to it. The set $\mathcal{K}+\mathcal{L}$ is a part of space lying between two planes parallel to $\mathcal{K}$ (Fig. 10).

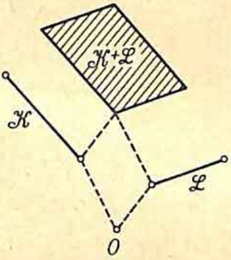


Fig. 9

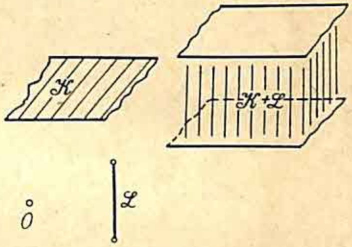


Fig. 10

4. $\mathcal{K}$ and $\mathcal{L}$ are circles of radii r_1 and r_2 with centres P_1 and P_2 (respectively) lying in the same plane π. Then $\mathcal{K}+\mathcal{L}$ is a circle of radius r_1+r_2 with the centre at the point P_1+P_2 lying in a plane parallel to π (Fig. 11).

2°. *The visualization of equations and inequalities of the first degree in two or three unknowns.* Consider a first-degree equation in two unknowns x and y:

$$ax + by + c = 0 \tag{1}$$

Interpreting x and y as coordinates of a point in the plane, it is natural to ask the question: What set is formed in the plane by the points whose coordinates satisfy equation (1), or in short what set of points is given by equation (1)?

We shall give the answer though the reader may already know it: *the set of points given by equation* (1) *is a straight line in the plane.* Indeed, if $b \neq 0$, then equation (1) is reduced to the form

$$y = kx + p$$

and this equation is known to give a straight line. If, however, $b = 0$, then the equation is reduced to the form

$$x = h$$

and gives a straight line parallel to the axis of ordinates.

A similar question arises concerning the inequality

$$ax + by + c \geqslant 0 \tag{2}$$

What set of points in the plane is given by inequality (2)?

Here again the answer is very simple. If $b \neq 0$, then the inequality is reduced to one of the following forms

$$y \geqslant kx + p \quad \text{or} \quad y \leqslant kx + p$$

It is easy to see that the first of these inequalities is satisfied by

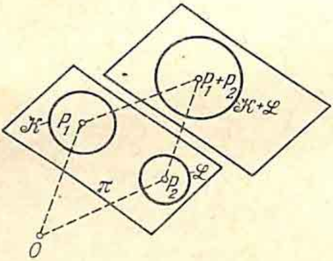


Fig. 11

all points lying "above" or on the straight line $y = kx + p$ and the second by all points lying "below" or on the line (Fig. 12). If, however, $b = 0$, then the inequality is reduced to one of the following forms

$$x \geqslant h \quad \text{or} \quad x \leqslant h$$

the first of them being satisfied by all points lying to the "right" of or on the straight line $x = h$ and the second by all points to the "left" of or on the line (Fig. 13).

Thus *equation* (1) *gives a straight line in the coordinate plane and inequality* (2) *gives one of the two half-planes into which this line divides the whole plane* (the line itself is considered to belong to either of these two half-planes).

We now want to solve similar questions with regard to the equation

$$ax + by + cz + d = 0 \tag{3}$$

and the inequality

$$ax + by + cz + d \geqslant 0 \tag{4}$$

of course, here x, y, z are interpreted as coordinates of a point *in space*. It is not difficult to foresee that the following result will be obtained.

Theorem. *Equation* (3) *gives a plane in space and inequality* (4) *gives one of the two half-spaces into which this plane divides the whole space* (the plane itself is considered to belong to one of these two half-spaces).

Proof. Of the three numbers a, b, c at least one is different from zero; let $c \neq 0$, for example. Then equation (3) is reduced to the form

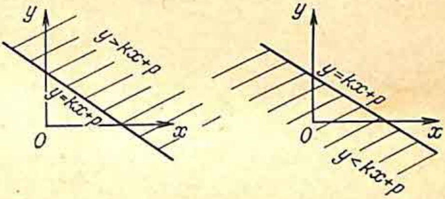


Fig. 12

$$z = kx + ly + p \tag{5}$$

Denote by $\mathscr{L}$ the set of all points M (x, y, z) which satisfy (5). Our aim is to show that $\mathscr{L}$ is a plane.

Find what points in $\mathscr{L}$ belong to the yOz coordinate plane. To do this, set $x = 0$ in (5) to obtain

$$z = ly + p \tag{6}$$

Thus the intersection of $\mathscr{L}$ with the yOz plane is the straight line u given in the plane by equation (6) (Fig. 14).

Similarly, we shall find that the intersection of $\mathscr{L}$ with the xOz plane is the straight line v given in the plane by the equation

$$z = kx + p \tag{7}$$

Both lines u and v pass through the point $P(0, 0, p)$.

Denote by π the plane containing the lines u and v. Show that π *belongs* to the set $\mathscr{L}$.

In order to do this it is sufficient to establish the following fact, viz. that a straight line passing through any point $A \in v$ and parallel to u belongs to $\mathscr{L}$.

First find a point B such that $OB \| u$. The equation $z = ly + p$ gives the straight line u in the yOz plane; hence the equation $z = ly$ gives a straight line parallel to u and passing through the origin (it is shown as dotted line in Fig. 14). We can take as B the point with the coordinates $y = 1$, $z = l$ which lies on this line.

An arbitrary point $A \in v$ has the coordinates x, 0, $kx + p$. The point B we have chosen has the coordinates 0, 1, l. The straight line passing through A and parallel to u consists of the points

$$A + sB = (x,\ 0,\ kx + p) + s(0,\ 1,\ l) =$$
$$= (x,\ s,\ kx + p + sl)$$

where s is an arbitrary number (see proposition (3) of section 1°).

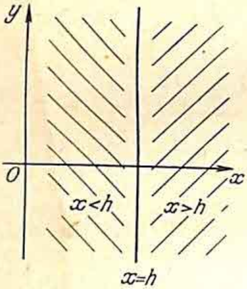


Fig. 13

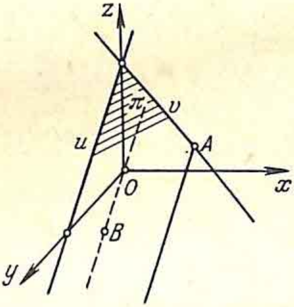


Fig. 14

It is easy to check that the coordinates of a point $A + sB$ satisfy equation (5), i.e. that $A + sB \in \mathscr{L}$. This proves that the plane π belongs wholly to the set $\mathscr{L}$.

It remains to make the last step, to show that $\mathscr{L}$ *coincides* with π or, in other words, that the set $\mathscr{L}$ does not contain any points outside π.

To do this, consider three points: a point M $(x_0,\ y_0,\ z_0)$ lying in the plane π, a point M' $(x_0,\ y_0,\ z_0 + \varepsilon)$ lying "above" the plane π $(\varepsilon > 0)$, and a point M'' $(x_0,\ y_0,\ z_0 - \varepsilon)$ lying "below" π (Fig. 15). Since $M \in \pi$, we have $z_0 = kx_0 + ly_0 + p$ and hence

$$z_0 + \varepsilon > kx_0 + ly_0 + p$$
$$z_0 - \varepsilon < kx_0 + ly_0 + p$$

This shows that the coordinates of the point M' satisfy the strict inequality

$$z > kx + ly + p$$

and the coordinates of the point M'' satisfy the strict inequality

$$z < kx + ly + p$$

Thereby M' and M'' do not belong to $\mathscr{L}$. This proves that $\mathscr{L}$ coincides with the plane π. In addition, it follows from our arguments that the set of all points satisfying the inequality

$$ax + by + cz + d \geqslant 0$$

is one of the two half-planes (either the "upper" or the "lower" one) into which the plane π divides the whole space.

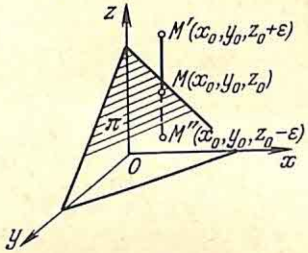


Fig. 15

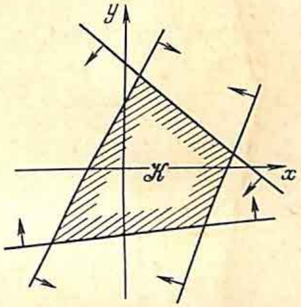


Fig. 16

2. Visualization of Systems of Linear Inequalities in Two or Three Unknowns

Let the following system of inequalities in two unknowns x and y be given

$$\left.\begin{array}{l} a_1x + b_1y + c_1 \geqslant 0 \\ a_2x + b_2y + c_2 \geqslant 0 \\ \cdots\cdots\cdots\cdots \\ a_mx + b_my + c_m \geqslant 0 \end{array}\right\} \qquad (1)$$

In the xOy coordinate plane the first inequality determines a half-plane Π_1, the second a half-plane Π_2, etc. If a pair of numbers x, y satisfies all the inequalities (1), then the corresponding point $M(x, y)$ belongs to all half-planes Π_1, Π_2 ..., Π_m simultaneously. In other words, the point M belongs to the *intersection* (common part) of the indicated half-planes. It is easy to see that the intersection of a finite number of half-planes is a polygonal region $\mathscr{K}$. Figure 16 shows one of the possible regions $\mathscr{K}$. The area of the region is shaded along the boundary. The inward direction of the strokes serves to indicate on which side of the given straight line the corresponding half-plane lies; the same is also indicated by the arrows.

The region $\mathscr{K}$ is called the *feasible region of the system* (1). Note from the outset that a feasible region is not always bounded: as a result of intersection of several half-planes an unbounded region may arise, as that of Fig. 17, for example. Having in mind the fact that the boundary of a region $\mathscr{K}$ consists of line segments (or whole straight lines) we say that $\mathscr{K}$ is a *polygonal feasible region of the system* (1) (we remark here that when a region $\mathscr{K}$ is bounded it is simply called a polygon*). Of course a case is possible where there is not a single point simultaneously belonging to all half-planes under consideration, i.e. where the region $\mathscr{K}$ is "empty"; this means that the system (1) is incompatible. One such case is presented in Fig. 18.

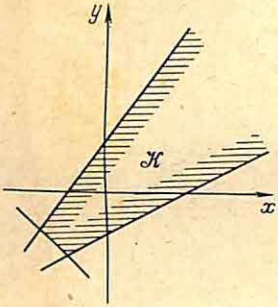


Fig. 17

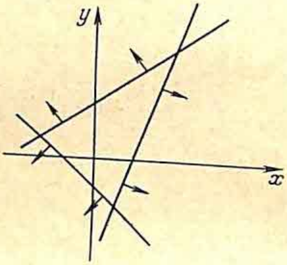


Fig. 18

Every feasible region $\mathscr{K}$ is convex. It will be recalled that according to the general definition *a set of points* (*in the plane or in space*) *is said to be convex if together with any two points A and B it contains the whole segment AB*. Figure 19 illustrates the difference between a convex and a nonconvex set. The convexity of a feasible region $\mathscr{K}$ ensues from the very way in which it is formed; for it is formed by intersection of several half-planes, and each half-plane is a convex set.

Lest there be any doubt, however, with regard to the convexity of $\mathscr{K}$, we shall prove the following lemma.

---

[*] Here to avoid any misunderstanding we must make a reservation. The word "polygon" is understood in the school geometry course to designate a closed *line* consisting of line segments, whereas in the literature on linear inequalities this term does not designate the line itself but *all the points of a plane which are spanned by it* (i. e. lie inside or on the line itself). It is in the sense accepted in the literature that the term "polygon" will be understood in what follows.

Lemma. *The intersection of any number of convex sets is a convex set.*

Proof. Let $\mathscr{K}_1$ and $\mathscr{K}_2$ be two convex sets and let $\mathscr{K}$ be their intersection. Consider any two points A and B belonging to $\mathscr{K}$ (Fig. 20). Since $A \in \mathscr{K}_1$, $B \in \mathscr{K}_1$ and the set $\mathscr{K}_1$ is convex, the segment AB belongs to $\mathscr{K}_1$. Similarly, the segment AB belongs to $\mathscr{K}_2$. Thus the segment AB simultaneously belongs to both sets $\mathscr{K}_1$ and $\mathscr{K}_2$ and hence to their intersection $\mathscr{K}$. This proves that $\mathscr{K}$ is a convex set. Similar arguments show that the intersection of any number (not necessarily two) of convex sets is a convex set.

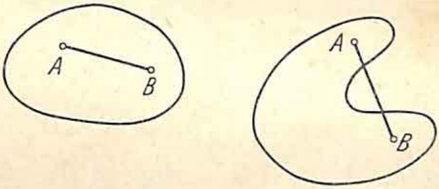


Fig. 19

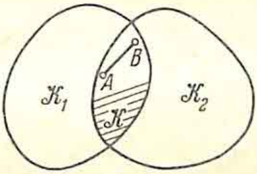


Fig. 20

Thus *the locus of the points whose coordinates satisfy all the inequalities* (1), *or equivalently the feasible region of the system* (1), *is a convex polygonal region* $\mathscr{K}$. *It is a result of intersection of all half-planes corresponding to the inequalities of the given system.*

Let us turn to the case involving three unknowns. Now we are given the system

$$\left.\begin{aligned} a_1x + b_1y + c_1z + d_1 &\geqslant 0 \\ a_2x + b_2y + c_2z + d_2 &\geqslant 0 \\ \dots\dots\dots\dots\dots&\dots \\ a_mx + b_my + c_mz + d_m &\geqslant 0 \end{aligned}\right\} \qquad (2)$$

As we know from Section 1, each of the above inequalities gives a half-plane. The region determined by the given system

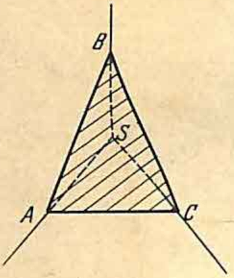


Fig. 21

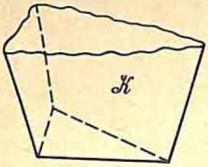


Fig. 22

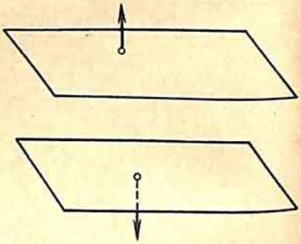


Fig. 23

will therefore be the intersection (common part) of m half-planes, and the intersection of a finite number of half-planes is a convex polyhedral region $\mathscr{K}$. Figure 21 exemplifies such a region for $m = 4$. In this example the region $\mathscr{K}$ is an ordinary tetrahedron (more strictly, $\mathscr{K}$ consists of all points lying inside and on the boundary of the tetrahedron); and in general it is not difficult to see that any convex polyhedron can be obtained as a result of the intersection of a finite number of half-planes.* Of course, a case is also possible where the region $\mathscr{K}$ is unbounded (where it extends into infinity); an example of such a region is represented in Fig. 22. Finally, it may happen that there are no points at all which satisfy all the inequalities under consideration (the system (2) is incompatible); then the region $\mathscr{K}$ is empty. Such a case is represented in Fig. 23.

---

* Here we must give an explanation of the kind given in the footnot on page 18. The thing is that in the school geometry course "polyhedron" refers to a *surface* composed of faces. We shall understand this term in a broader sense, i.e. as referring to *the set of all points of space spanned by the surface* rather than to the surface itself, the set of course including the surface itself but only as its part.

Particular attention should be given to the case where the system (2) contains among others the following two inequalities:

$$ax + by + cz + d \geqslant 0$$
$$-ax - by - cz - d \geqslant 0$$

which can be replaced by the sole equation

$$ax + by + cz + d = 0$$

The latter gives a plane π in space. The remaining inequalities (2) will separate in the plane π a convex polygonal region which will be the feasible region of the system (2). It is seen that *a particular case of a convex polygonal region in space may be represented by a convex polygonal region in the plane*. In Fig. 24 the region $\mathscr{K}$ is a triangle formed by the intersection of five half-planes, two of them being bounded by the "horizontal" plane π and the remaining three forming a "vertical" trihedral prism.

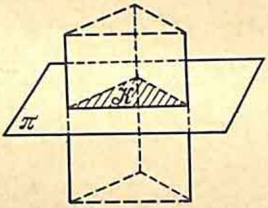


Fig. 24

By analogy with the case involving two unknowns, we call the region $\mathscr{K}$ the *feasible region of the system* (2). We shall emphasize once again the fact that a region $\mathscr{K}$ being the intersection of a number of half-planes is necessarily convex.

Thus *the system* (2) *gives a convex polyhedral region* $\mathscr{K}$ *in space. This region results from intersection of all half-planes corresponding to the inequalities of the given system.*

If the region $\mathscr{K}$ is bounded, it is simply called the *feasible polyhedron of the system* (1).

25.8.94
8714

3. The Convex Hull of a System of Points

Imagine a plane in the shape of an infinite sheet of plywood to have pegs driven into it at points A_1, A_2, ..., A_p. Having made a rubber loop, stretch it well and let it span all the pegs (see the dotted line in Fig. 25). Then allow it to tighten, to the extent permitted by the pegs of course. The set of the points spanned by the loop after the tightening is represented in Fig. 25 by the shaded area. It appears to be a convex polygon. The latter is called the *convex hull of the system of the points* A_1, A_2, ..., A_p.

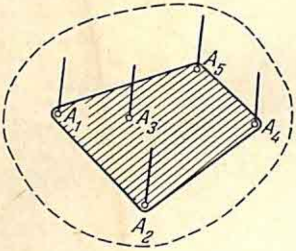


Fig. 25

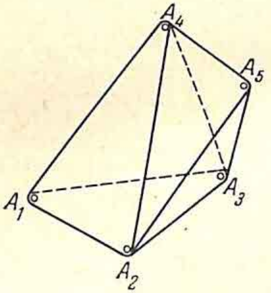


Fig. 26

If the points A_1, A_2, ..., A_p are located in space rather than in the plane, then a similar experiment is rather hard to put into practice. Let us give rein to our imagination, however, and suppose that we have managed to confine the points A_1, A_2, ..., A_p in a bag made of tight rubber film. Left to its own devices, the bag will tighten until prevented from doing so by some of the points. Finally, a time will arrive when any further tightening is no longer possible (Fig. 26). It is fairly clear that by that time the bag will have taken shape of a convex polyhedron with vertices at some of the points $A_1, A_2, ..., A_p$. The region of space spanned by this polyhedron is again called the convex hull of the system of the points A_1, A_2, ..., A_p.

Very visual as it is, the above definition of the convex hull is not quite faultless from the standpoint of "mathematical strictness". We now define that notion strictly.

Let A_1, A_2, ..., A_p be an arbitrary set of points (in the plane or in space). Consider all sorts of points of the form

$$s_1A_1 + s_2A_2 + ... + s_pA_p \tag{1}$$

where $s_1, s_2, \dots, s_p$ are any nonnegative numbers whose sum is one:

$$s_1,\ s_2,\ \dots,\ s_p \geqslant 0 \quad \text{and} \quad s_1 + s_2 + \dots + s_p = 1 \tag{2}$$

Definition. *The set of points of the form* (1) *with condition* (2) *is called the convex hull of the system of the points* $A_1, A_2, \dots, A_p$ *and denoted by*

$$\langle A_1,\ A_2,\ \dots,\ A_p \rangle$$

To make sure that this definition does not diverge from the former, first consider the cases $p = 2$ and $p = 3$. If $p = 2$, then we are given two points A_1 and A_2. The set $\langle A_1, A_2 \rangle$, as indicated by proposition (1) of section 1, is a segment A_1A_2.

If $p = 3$, then we are given three points A_1, A_2 and A_3. We show that the set $\langle A_1, A_2, A_3 \rangle$ consists of all the points lying inside and on the sides of the triangle $A_1A_2A_3$.

Moreover, we prove the following lemma.

Lemma. *The set* $\langle A_1, \dots, A_{p-1}, A_p \rangle$ *consists of all segments joining the point* A_p *with the points of the set* $\langle A_1, \dots, A_{p-1} \rangle$.

Proof. For notational convenience, denote the set $\langle A_1, \dots, A_{p-1} \rangle$ by $\mathcal{M}_{p-1}$ and the set $\langle A_1, \dots, A_{p-1}, A_p \rangle$ by $\mathcal{M}_p$.

Consider any point $A \in \mathcal{M}_p$. It is of the form

$$A = s_1A_1 + \dots + s_{p-1}A_{p-1} + s_pA_p$$

where

$$s_1,\ \dots,\ s_p \geqslant 0, \quad s_1 + \dots + s_p = 1$$

If $s_p = 0$, then $A \in \mathcal{M}_{p-1}$; thus the set $\mathcal{M}_{p-1}$ is a part of $\mathcal{M}_p$. If $s_p = 1$, then $A = A_p$; thus the point A_p belongs to $\mathcal{M}_p$. So $\mathcal{M}_p$ contains $\mathcal{M}_{p-1}$ and the point A_p. We now show that any segment $A'A_p$ where $A' \in \mathcal{M}_{p-1}$ belongs wholly to $\mathcal{M}_p$.

If A is a point of such a segment, then

$$A = tA' + sA_p \quad (t,\ s \geqslant 0,\ t + s = 1)$$

On the other hand, by the definition of the point A' we have

$$A' = t_1A_1 + \dots + t_{p-1}A_{p-1}$$

$$(t_1,\ \dots,\ t_{p-1} \geqslant 0, \quad t_1 + \dots + t_{p-1} = 1)$$

hence

$$A = tt_1A_1 + \dots + tt_{p-1}A_{p-1} + sA_p$$

Setting $tt_1 = s_1, \dots, tt_{p-1} = s_{p-1}, s = s_p$, we have (1), (2). This proves that $A \in \mathcal{M}_p$. So any of the above segments belongs wholly to $\mathcal{M}_p$

Now it remains for us to check that the set $\mathcal{M}_p$ does not contain anything but these segments, i.e. that *any point of $\mathcal{M}_p$ belongs to one of the segments under consideration.*

Let $A \in \mathcal{M}_p$. Then we have (1), (2). It can be considered that $s_p \neq 1$, otherwise $A = A_p$ and there is nothing to be proved. But if $s_p \neq 1$, then $s_1 + \dots + s_{p-1} = 1 - s_p > 0$, therefore we can write down

$$A = (s_1 + \dots + s_{p-1}) \left[\frac{s_1}{s_1 + \dots + s_{p-1}} A_1 + \dots \right.$$

$$\left. \dots + \frac{s_{p-1}}{s_1 + \dots + s_{p-1}} A_{p-1} \right] + s_p A_p$$

The expression in square brackets determines some point A' belonging to $\mathcal{M}_{p-1}$, for the coefficients of $A_1, \dots, A_{p-1}$ are nonnegative in this expression and their sum is one. So

$$A = (s_1 + \dots + s_{p-1}) A' + s_p A_p$$

Since the coefficients of A' and A_p are also nonnegative and their sum is one, the point A lies on the segment $A'A_p$. This completes the proof of the lemma.

Now it is not difficult to see that the visual definition of the convex hull given at the beginning of this section and the strict definition which follows it are equivalent. Indeed, whichever of the two definitions of the convex hull may be assumed as the basis, in either case going over from the convex hull of the system $A_1, \dots, A_{p-1}$ to that of the system $A_1, \dots, A_{p-1}, A_p$ follows one and the same rule, namely the point A_p must be joined by segments to all the points of the convex hull for $A_1, \dots, A_{p-1}$ (this rule is immediately apparent when the convex hull is visually defined and in the strict definition it makes the content of the lemma). If we now take into account the fact that according to both definitions we have for $p = 2$ one and the same set, the segment A_1A_2, the equivalence of both definitions becomes apparent.

The term "convex hull" has not yet been quite justified by us, however, for we have not yet shown that the *set $\langle A_1, A_2, \dots, A_p \rangle$ is always convex*. We shall do it now.

Let A and B be two arbitrary points of this set:

$$A = s_1 A_1 + s_2 A_2 + \dots + s_p A_p$$
$$B = t_1 A_1 + t_2 A_2 + \dots + t_p A_p$$

where

$$s_1, \dots, s_p, t_1, \dots, t_p \geqslant 0$$
$$s_1 + \dots + s_p = t_1 + \dots + t_p = 1 \tag{3}$$

Any point C of the segment AB is of the form

$$C = sA + tB$$

where

$$s,\ t \geqslant 0, \quad s + t = 1 \tag{4}$$

which gives

$$C = s(s_1 A_1 + \ldots + s_p A_p) + t(t_1 A_1 + \ldots + t_p A_p) =$$
$$= (ss_1 + tt_1) A_1 + \ldots + (ss_p + tt_p) A_p$$

The numbers preceding $A_1, \ldots, A_p$ as coefficients are nonnegative and their sum amounts to one (which follows from (3), (4)). This means that the point C belongs to the set $\langle A_1, A_2, \ldots, A_p \rangle$, i.e. this set is convex.

At the same time it is easy to see that $\langle A_1, A_2, \ldots, A_p \rangle$ is the *least of all convex sets that contain the starting points* $A_1, A_2, \ldots, A_p$, viz. that it is contained in any of these sets. This statement follows directly from the lemma proved above and from the definition of a convex set.

This explains the name "convex hull" and at the same time provides yet another explanation of the fact why the set $\langle A_1, A_2, \ldots, A_p \rangle$ can be obtained using the method described at the beginning of the section. For the set spanned by a rubber loop (or film) after the latter has tightened to its limit around the system of the points $A_1, A_2, \ldots, A_p$ is exactly the least convex set that contains the indicated points.

4. A Convex Polyhedral Cone

Let's begin with a definition.

A convex polyhedral cone is the intersection of a finite number of half-spaces whose boundary surfaces pass through a common point; the latter being called the vertex of the cone.

We shall first of all point out how the notion of a convex polyhedral cone is related to systems of linear inequalities. We shall confine ourselves to a particular case, namely a case where the vertex of a cone is the origin of coordinates. This means that all boundary planes pass through the origin, and the equation of a plane passing through the origin is of the form

$$ax + by + cz = 0$$

(the absolute term in the equation must be equal to zero, otherwise (0, 0, 0) will not satisfy the equation). Thus *a convex po-*

lyhedral cone with the vertex at the origin is the feasible region of a system of homogeneous inequalities:

$$\left.\begin{array}{l} a_1x + b_1y + c_1z \geqslant 0 \\ a_2x + b_2y + c_2z \geqslant 0 \\ \dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz \geqslant 0 \end{array}\right\}$$

The reverse is true of course; the feasible region of a homogeneous system of inequalities is always a convex polyhedral cone with the vertex at the origin.

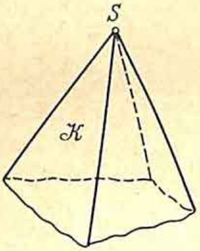


Fig. 27

An example of a convex polyhedral cone may be provided by a convex region in space whose boundary is a polyhedral angle with the vertex S, a kind of infinite convex pyramid without a base and extending unboundedly from the vertex (Fig. 27 represents one of such pyramids which has four faces). Other, less interesting cases are possible; for example:

1. A half-space (Fig. 28, *a*). In such a "cone" the role of the vertex may be played by any point $S \in \pi$, where π is the boundary plane of a given half-space.

2. The intersection of two half-spaces whose boundary planes intersect along a straight line l (Fig. 28, *b*). The role of the vertex may be played by any point $S \in l$.

3. A plane. It is clear that any plane π in space can be considered as the intersection of two half-spaces lying on different sides of π (Fig. 28, *c*). In this case the role of the vertex may be played by any point $S \in \pi$.

4. A half-plane (Fig. 28, *d*). The vertex S is any point of the boundary line.

5. A straight line. Every line l in space can be obtained by intersection of three half-spaces whose boundary planes pass through l (Fig. 28, *e*). The vertex S is any point of the line l.

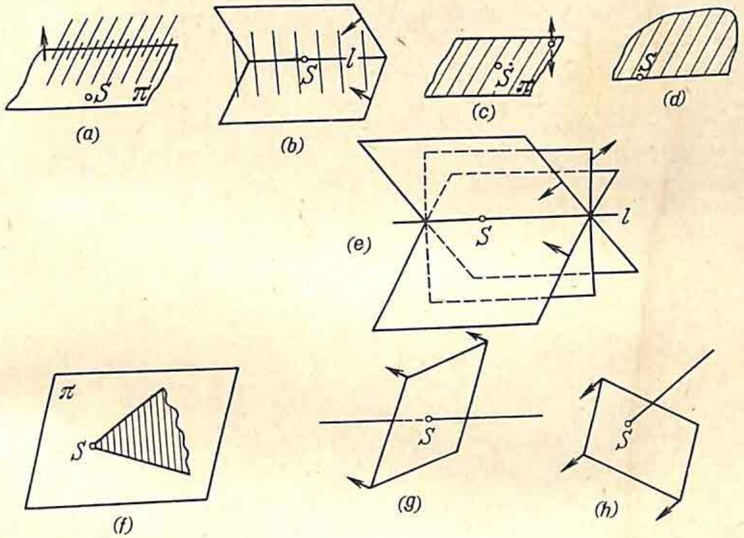


Fig. 28

6. An angle (less than 180°) in an arbitrary plane π (Fig. 28, *f*). One can obtain an angle by intersecting the plane π with two half-spaces (precisely how?).

7. A ray (Fig. 28, *g*). A ray can be considered as the intersection of a straight line and a half-space. The vertex S is the beginning of the ray.

8. A point. This "cone" can be obtained by taking the common part of a ray and the corresponding half-space (Fig. 28, *h*).

Of course, the enumerated examples 1-8 diverge (to a greater or lesser extent) from the usage of the word "cone", but we are compelled to reconcile ourselves to this if we are to preserve the general definition of a convex polyhedral cone given at the beginning of this section.

We now try to show in a few words that *the sets pointed out above exhaust all polyhedral convex cones in space.*

Let p denote the number of half-spaces whose intersection is the cone $\mathscr{K}$ under consideration. If $p = 1$, then our statement is valid, for then $\mathscr{K}$ is a half-space. A simple argument which is left to the reader shows that, if our statement is valid for a cone which results from intersection of p half-spaces, it is also valid for a cone formed by intersection of $p + 1$ half-spaces. It follows according to the principle of complete mathematical induction that our statement is valid for any p.

Convex polyhedral cones possess many interesting properties. It is beyond the scope of the present book to go into these subjects, so we shall confine ourselves to a few simplest propositions.

We introduce one more definition or notation, if you please.

Let $B_1, B_2, \ldots, B_q$ be an arbitrary set of a finite number of points (in space). *The symbol* $(B_1, B_2, \ldots, B_q)$ *will denote a set of points of the form*

$$t_1B_1 + t_2B_2 + \ldots + t_qB_q$$

where $t_1, t_2, \ldots, t_q$ *are arbitrary nonnegative numbers.*

What is the geometric meaning of the set $(B_1, B_2, \ldots, B_q)$? It is clear from the definition that it is the sum of the sets $(B_1), (B_2), \ldots, (B_q)$; we must first see, therefore, what is the geometric meaning of the set (B), i.e. of the set of points of the form tB, where t is any nonnegative number and B a fixed point. But the answer to the last question is obvious: if B is the origin, then the set (B) also coincides with the origin; otherwise (B) is a ray emerging from the origin and passing through the point B. Now we remark that the sum of any set and the origin is again the same set; hence it is clear that when studying the sets $(B_1, B_2, \ldots, B_q)$ we shall lose nothing if *we consider all the points* $B_1, B_2, \ldots, B_q$ *to be different from the origin. Then the set* $(B_1, B_2, \ldots, B_q)$ *will be the sum of the rays* $(B_1), (B_2), \ldots, (B_q)$.

The last remark makes the following lemma almost obvious.

Lemma. *The set* $(B_1, \ldots, B_{q-1}, B_q)$ *is a union of segments joining each point of the set* $(B_1, \ldots, B_{q-1})$ *with each point of the ray* (B_q).

The strict proof of the lemma is carried out according to the same plan as the proof of the similar lemma of Section 3; the reader is advised to carry it out independently.

It is easily deduced from the lemma that (B_1, B_2) is an angle, a straight line or a ray (Fig. 29, *a*, *b*, *c*). It is then readily established that (B_1, B_2, B_3) is one of the following sets: an infinite trihedral pyramid, a plane, a half-plane, an angle, a straight line or a ray. Now it becomes clear that there must exist a close relation between the sets (B_1), (B_2), ..., (B_q) and convex polyhedral cones. Such a relation does in fact exist. For greater intelligibility we shall formulate the corresponding propositions as two theorems.

Theorem 1. *The set* $(B_1, B_2, ..., B_q)$ *either coincides with the whole of space or else is a convex polyhedral cone with vertex at the origin.*

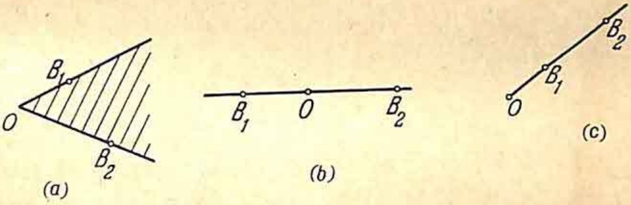


Fig. 29

That the set $(B_1, B_2, ..., B_q)$ can indeed coincide with the whole of space is shown by the following example. Consider four points B_1, B_2, B_3, B_4 lying in such a way that the rays (B_1), (B_2), (B_3), (B_4) form pairwise obtuse angles (Fig. 30). Each of the sets (B_1, B_2, B_3), (B_1, B_2, B_4), (B_1, B_3, B_4), (B_2, B_3, B_4) is an infinite trihedral pyramid with vertex at the origin. The set (B_1, B_2, B_3, B_4) apparently contains all of these pyramids; and the union of the pyramids coincides with the whole of space!

Theorem 2. *Any convex polyhedral cone with vertex at the origin is a set of the form* $(B_1, B_2, ..., B_q)$.

The proof of Theorem 1 will be carried out on general lines. We shall take advantage of the method of complete mathematical induction. The statement of the theorem for $q = 1$ is obvious. Now we suppose that the theorem is valid for sets of the form $(B_1, ..., B_q)$ and, guided by this fact, prove it to be valid for the sets $(B_1, ..., B_q, B_{q+1})$.

According to the induction hypothesis, $(B_1, ..., B_q)$ is either the whole of space or a convex polyhedral cone in it. As to the first case there is as a matter of fact nothing to be proved in it, for then $(B_1, ..., B_q, B_{q+1})$ is also the whole of space. Let the second case occur and $(B_1, ..., B_q)$ be a convex polyhedral cone $\mathscr{K}$. According to the lemma, the set $(B_1, ..., B_q, B_{q+1})$ is a union of segments joining each point of the set $\mathscr{K}$ with each point of the ray (B_{q+1}); and, as shown earlier, any convex polyhedral cone $\mathcal{K}$ is either an infinite convex pyramid or one of the sets 1 to 8. Having considered the above union of segments for each of these cases, it is not difficult to make sure (check it for yourself) that either it coincides with the whole of space or is again a convex polyhedral cone. Thus the theorem is true for sets of the kind (B_1) and for $(B_1, \ldots, B_q, B_{q+1})$ as well, as soon as we suppose it to be valid for $(B_1, \ldots, B_q)$. Hence it follows that the theorem is true for any q.

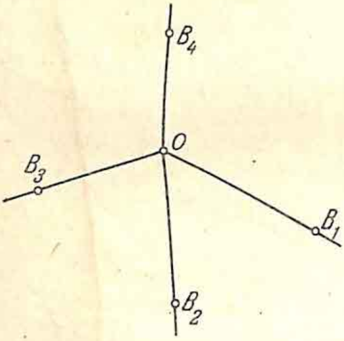


Fig. 30

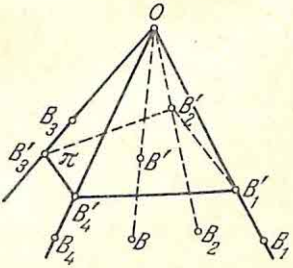


Fig. 31

Proof of Theorem 2. Let $\mathcal{K}$ be a convex polyhedral cone with vertex at the origin O. As already stated, $\mathcal{K}$ is either an infinite convex pyramid or one of the sets 1 to 8.

Let $\mathcal{K}$ be a pyramid. We choose a point on each of its edges to get a system of points $B_1, B_2, \ldots B_q$. We state that the *set* $(B_1, B_2, \ldots, B_q)$ *is exactly* $\mathcal{K}$.

To prove this, consider a plane π intersecting all the edges of the pyramid $\mathcal{K}$. We get the points $B_1', B_2', \ldots, B_q'$ (Fig. 31). Apparently,

$$B_1' = k_1 B_1, \quad B_2' = k_2 B_2, \quad B_q' = k_q B_q \tag{1}$$

where $k_1, k_2, \ldots, k_q$ are nonnegative numbers.

Now suppose B is a point of the pyramid different from the vertex O. The ray OB intersects with the plane π at a point B'. It is obvious that B' belongs to the convex hull of the system $B_1', B_2', \ldots, B_q'$ and hence

$$B' = s_1 B_1' + s_2 B_2' + \ldots + s_q B_q'$$

where $s_1, s_2, \ldots, s_q$ are nonnegative numbers whose sum is one.

Now taking into account (1) we have

$$B' = s_1k_1B_1 + s_2k_2B_2 + \dots + s_qk_qB_q$$

and if in addition it is remembered that $B' = kB$ $(k > 0)$, we find that

$$B = t_1B_1 + t_2B_2 + \dots + t_qB_q$$

where

$$t_i = \frac{s_ik_i}{k} \quad (i = 1, 2, \dots, q)$$

Thus we have shown that any point B of the pyramid $\mathscr{K}$ belongs to the set $(B_1, B_2, \dots, B_q)$. The converse (i.e. that any point of the set $(B_1, B_2, \dots, B_q)$ belongs to $\mathscr{K}$) is obvious. So $\mathscr{K}$ coincides with $(B_1, B_2, \dots, B_q)$.

The case where $\mathscr{K}$ is one of the exceptional sets 1 to 8 can be proved without much trouble and is left to the reader.

5. The Feasible Region of a System of Linear Inequalities in Two Unknowns

Our task now is to give an effective description of all solutions of a system of linear inequalities. In the present section we shall deal with systems involving two unknowns x and y. In spite of the fact that the number of unknowns is not large (only two), we shall try to carry out the analysis of these systems from general positions, i.e. so that the results obtained may be easily extended to systems in a larger number of unknowns.

The solution of any system of linear inequalities is in the long run reduced to the solution of a *number of systems of linear equations*. We shall regard the solution of a system of linear equations as something simple, as an elementary operation, and shall not be confused if, to realize the proposed method, we have to perform this operation many times.

1°. *The necessary lemmas*. Given a system of inequalities

$$\left.\begin{array}{l} a_1x + b_1y + c_1 \geqslant 0 \\ a_2x + b_2y + c_2 \geqslant 0 \\ \dots\dots\dots\dots \\ a_mx + b_my + c_m \geqslant 0 \end{array}\right\} \qquad (1)$$

It is found advisable to consider side by side with it the corresponding system of *homogeneous inequalities*

$$\left.\begin{array}{l} a_1x + b_1y \geqslant 0 \\ a_2x + b_2y \geqslant 0 \\ \dots\dots\dots \\ a_mx + b_my \geqslant 0 \end{array}\right\} \qquad (2)$$

as well as the corresponding system of *homogeneous equations*

$$\left.\begin{array}{l} a_1x + b_1y = 0 \\ a_2x + b_2y = 0 \\ \dots\dots\dots \\ a_mx + b_my = 0 \end{array}\right\} \qquad (3)$$

We denote the feasible region of the system (1) in the xOy coordinate plane by $\mathscr{K}$, that of the system (2) by $\mathscr{K}_0$ and that of the system (3) by $\mathscr{L}$. Obviously, $\mathscr{L} \subset \mathscr{K}_0$, where the symbol $\subset$ stands for the words "is contained in"*.

Lemma 1. *The following inclusion holds*

$$\mathscr{K} + \mathscr{K}_0 \subset \mathscr{K}$$

i.e. the sum of any solution of a given system of inequalities and any solution of the corresponding homogeneous system of inequalities is again a solution of the given system.

Proof. Let A be an arbitrary point of $\mathscr{K}$ and let B be an arbitrary point of $\mathscr{K}_0$. Then the following inequalities are valid

$$\begin{array}{lcl} a_1x_A + b_1y_A + c_1 \geqslant 0 & & a_1x_B + b_1y_B \geqslant 0 \\ a_2x_A + b_2y_A + c_2 \geqslant 0 & \text{and} & a_2x_B + b_2y_B \geqslant 0 \\ \dots\dots\dots\dots & & \dots\dots\dots \\ a_mx_A + b_my_A + c_m \geqslant 0 & & a_mx_B + b_my_B \geqslant 0 \end{array}$$

Adding each inequality written on the left to the corresponding inequality on the right we have

$$\begin{array}{l} a_1(x_A + x_B) + b_1(y_A + y_B) + c_1 \geqslant 0 \\ a_2(x_A + x_B) + b_2(y_A + y_B) + c_2 \geqslant 0 \\ \dots\dots\dots\dots\dots\dots \\ a_m(x_A + x_B) + b_m(y_A + y_B) + c_m \geqslant 0 \end{array}$$

---

* One should not confuse the symbol $\subset$ with the symbol $\in$ introduced earlier. The latter is used when speaking of the *belonging of a point to a set*. If, however, one wants to record the fact that *one set is a part of another*, the symbol $\subset$ is used.

These inequalities imply that the pair of numbers $x_A + x_B$, $y_A + y_B$, the coordinates of the point $A + B$, is a solution of the original system (1), i.e. that $A + B \in \mathscr{K}$. Thus the lemma is proved.

Lemma 2. (1) *If a ray with the beginning at the point A belongs wholly to the set* $\mathscr{K}$ *and P is an arbitrary point of the ray, then* $P - A \in \mathscr{K}_0$.

(2) *If a straight line belongs wholly to* $\mathscr{K}$ *and A, P are two arbitrary points of the line, then* $P - A \in \mathscr{L}$.

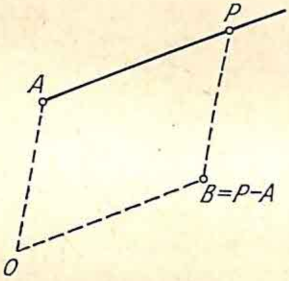


Fig. 32

Proof of (1). Denote the point $P - A$ by B. The ray under consideration consists of points of the form

$$A + sB \tag{4}$$

where s is an arbitrary nonnegative number (Fig. 32). Any of these points is, according to the premises, a solution of the system (1), i.e.

$$\left.\begin{array}{l} a_1(x_A + sx_B) + b_1(y_A + sy_B) + c_1 \geqslant 0 \\ a_2(x_A + sx_B) + b_2(y_A + sy_B) + c_2 \geqslant 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_m(x_A + sx_B) + b_m(y_A + sy_B) + c_m \geqslant 0 \end{array}\right\} \tag{5}$$

Consider, for example, the first of these inequalities. It can be written down in the following form

$$(a_1 x_A + b_1 y_A + c_1) + s(a_1 x_B + b_1 y_B) \geqslant 0$$

Since this inequality holds for *any* $s \geqslant 0$, the coefficient of s must clearly be a nonnegative number:

$$a_1 x_B + b_1 y_B \geqslant 0$$

Similarly, consideration of the remaining inequalities (5) leads to

$$\begin{gathered} a_2x_B + b_2y_B \geqslant 0 \\ \cdots\cdots\cdots\cdots \\ a_mx_B + b_my_B \geqslant 0 \end{gathered}$$

whence we see that the point B belongs to the set $\mathscr{K}_0$.

The proof of (2) is carried out in a similar way. The straight line under consideration consists of points of the form (4) where s is an arbitrary number. Therefore inequalities (5) are valid for any value of s. Hence it follows that in each of these inequalities the total coefficient of s must be equal to zero, i.e.

$$\begin{gathered} a_1x_B + b_1y_B = 0 \\ a_2x_B + b_2y_B = 0 \\ \cdots\cdots\cdots\cdots \\ a_mx_B + b_my_B = 0 \end{gathered}$$

Therefore $B \in \mathscr{L}$. Thus the lemma is proved.

It is easy to see that Lemmas 1 and 2 are valid for systems involving any number of unknowns.

2°. *The case where the system of inequalities* (1) *is normal*. Consider again the system of inequalities (1) and the corresponding system of homogeneous equations (3). The latter system has an obvious solution $x = 0$, $y = 0$ which is called a *zero* solution. In order to investigate the system (1) it turns out to be important to know if the system (3) has any nonzero solutions either. In view of this we introduce the following

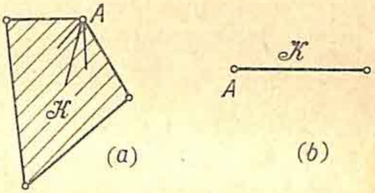
Definition. *A system of linear inequalities is said to be normal if the corresponding system of linear homogeneous equations has only a zero solution.*

In other words, a system of inequalities is normal if the set $\mathscr{L}$, the feasible region of the corresponding homogeneous system of equations, defined above contains only a single point (the origin of coordinates).

The concept of normal system is meaningful with any number of unknowns, of course.

It is not difficult to show that *a compatible system of inequalities is normal if and only if its feasible region $\mathscr{K}$ contains no straight lines.*

Indeed, if the system is normal, i.e. the set $\mathscr{L}$ contains only the origin of coordinates, then the region $\mathscr{K}$ does not contain any straight lines, which follows directly from the second statement of Lemma 2. If the system is not normal, then the set $\mathscr{L}$ contains at least one point B different from the origin of coordinates. Of course, all points of the form kB, where k is any number, also belong to $\mathscr{L}$*. But in this case, whatever the point $P \in \mathscr{K}$ (and such a point is sure to exist for the system is compatible and the region $\mathscr{K}$ is therefore not empty), the set of all points of the form $P + kB$ (where k is any number) belongs, according to Lemma 1, to $\mathscr{K}$. We know that the set is a straight line. So when a system is not normal, the region $\mathscr{K}$ contains a straight line. This completes the proof of the above proposition.



(a) (b)

Fig. 33

In this section we shall study the feasible region of the system (1) under the supposition that the system is compatible (the region $\mathscr{K}$ is not empty) and normal.

From the fact that the region $\mathscr{K}$ does not contain any straight lines it first of all follows that it must necessarily have *vertices*. The term "vertex" is understood by us in the following sense (close to the intuitive understanding of the word "vertex").

The vertex of a region $\mathscr{K}$ *is a point of the region which is not an interior point for any segment which lies wholly in* $\mathscr{K}$. In other words, the vertex is a point $A \in \mathscr{K}$ which has the following property: any segment belonging to $\mathscr{K}$ and passing through A must have its beginning or end at this point (Fig. 33, *a* and *b*, where the point A is one of the vertices; in Fig. 33, *b* the region $\mathscr{K}$ is a segment).

We now explain at greater length why the convex set $\mathscr{K}$ we are interested in has vertices. If $\mathscr{K}$ lies on a straight line, then it is either a separate point, a segment or a ray, and the existence of vertices is evident. If $\mathscr{K}$ does not lie on a straight line, however,

---

* If the numbers x, y, z, the coordinates of the point B, satisfy a homogeneous system of equations, then the numbers kx, ky, kz, the coordinates of the point kB, satisfy this system too.

consider the boundary of this set. It consists of segments and rays ($\mathcal{K}$ does not contain any whole lines). The end of any of these segments and the beginning of any of the rays will evidently be the vertices of $\mathcal{K}$.

One can find the vertices of the region $\mathcal{K}$ without much trouble. Notice first of all that in the xOy coordinate plane the ith inequality of the system (1) corresponds to the half-plane whose boundary line l_i is given by the equation

$$a_ix + b_iy + c_i = 0 \quad (i = 1, 2, ..., m)$$

Evidently, the point A of the region $\mathcal{K}$ is a vertex if and only if it belongs to two *different* boundary lines.

Let us agree to call *regular* any subsystem of two equations of the system

$$\left.\begin{array}{l} a_1x + b_1y + c_1 = 0 \\ a_2x + b_2y + c_2 = 0 \\ \cdots\cdots\cdots\cdots \\ a_mx + b_my + c_m = 0 \end{array}\right\} \tag{6}$$

provided it has a *unique* solution (x,y).

From the above description of the vertex now results the following method of finding the vertices of the region $\mathcal{K}$.

In order to find all the vertices one should find the solutions of all regular subsystems of the system (6) *and pick out those which satisfy the original system* (1).

Since the number of regular subsystems does not exceed C_m^2, the number of combinations of m things 2 at a time, the number of the vertices of the region $\mathcal{K}$ cannot be greater than that either. So the *number of the vertices is finite.*

Remark. It follows from the above that, if the region $\mathcal{K}$ of the solutions of a normal system has no vertex, it is empty—the system has no solutions (it is incompatible).

Example 1. Find all the vertices of the region $\mathcal{K}$ given by the system of inequalities

$$\left.\begin{array}{r} x + y + 1 \geqslant 0 \\ x - 2y - 2 \geqslant 0 \\ 2x - y - 4 \geqslant 0 \end{array}\right\}$$

On solving the subsystems

$$\left.\begin{array}{r} x + y + 1 = 0 \\ x - 2y - 2 = 0 \end{array}\right\} \quad \left.\begin{array}{r} x + y + 1 = 0 \\ 2x - y - 4 = 0 \end{array}\right\} \quad \left.\begin{array}{r} x - 2y - 2 = 0 \\ 2x - y - 4 = 0 \end{array}\right\}$$

(all of them prove regular) we find three points:

$$(0, -1), \quad (1, -2), \quad (2, 0)$$

of which only the second and the third satisfy all the inequalities given. So it is the points

$$A_1(1, -2) \quad \text{and} \quad A_2(2, 0)$$

that are the vertices of the region $\mathscr{K}$.

Let us return to the system (1). Let

$$A_1, A_2, \ldots, A_p$$

be all the vertices of the region $\mathscr{K}$. The set $\langle A_1, A_2, \ldots, A_p \rangle$, the convex hull of the system of points $A_1, A_2, \ldots, A_p$, also belongs to $\mathscr{K}$ (for $\mathscr{K}$ is a convex region!). But in that case, according to Lemma 1, the set

$$\langle A_1, A_2, \ldots, A_p \rangle + \mathscr{K}_0$$

belongs to $\mathscr{K}$ too. We shall show that as a matter of fact this sum coincides with $\mathscr{K}$, that we have the following

Theorem. *If the system of inequalities is normal, then*

$$\mathscr{K} = \langle A_1, A_2, \ldots, A_p \rangle + \mathscr{K}_0 \tag{7}$$

where $A_1, A_2, \ldots, A_p$ *are all the vertices of the region* $\mathscr{K}$.

Proof. Let P be an arbitrary point of the region $\mathscr{K}$ different from the vertices of the region. The line A_1P intersects the convex region $\mathscr{K}$ either along a segment A_1A (Fig. 34) or along a ray with origin at A_1 (Fig. 35). In the second case $P - A_1 \in \mathscr{K}_0$ (Lemma 2), hence $P \in A_1 + \mathscr{K}_0$. In the first case we reason as follows, however. If the point A lies on the bounded edge A_iA_j of the region $\mathscr{K}$ (as in Fig. 34), then P belongs to the convex hull of the points A_1, A_i, A_j; if, however, the point A lies on an unbounded edge with origin at vertex A_i (Fig. 36), then, according to Lemma 1, we have $A \in A_i + \mathscr{K}_0$ and thereby $P \in \langle A_1, A_i \rangle + \mathscr{K}_0$. Thus in all cases the point P proves to belong to the set $\langle A_1, A_2, \ldots, A_p \rangle + \mathscr{K}_0$. Thus the theorem is proved. Since we are already familiar with the method of finding vertices, the only thing we lack for a complete description of the region $\mathscr{K}$ is the ability to find the region $\mathscr{K}_0$. The latter is the feasible region of the homogeneous normal system (2) which we now proceed to describe.

3°. *The homogeneous normal system of inequalities* (2). Each of the inequalities (2) determines a half-plane whose boundary line passes through the origin of coordinates. It is the common part of these half-planes that is $\mathscr{K}_0$.

Among the boundary lines there are in this case at least two different ones (for the system (2) is normal!). Hence $\mathscr{K}_0$ either coincides with the origin of coordinates ($x = 0$, $y = 0$), or is a ray with vertex at the origin of coordinates or an angle smaller than 180° with vertex at origin. If we know two points B_1 and B_2 lying on different sides of the angle (Fig. 37), then all the

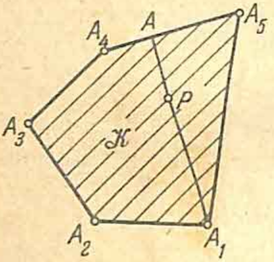


Fig. 34

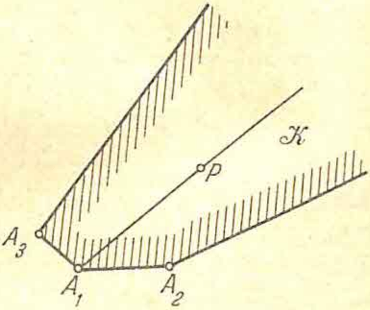


Fig. 35

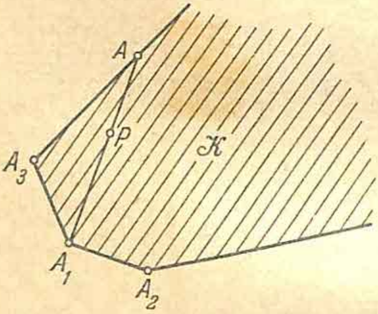


Fig. 36

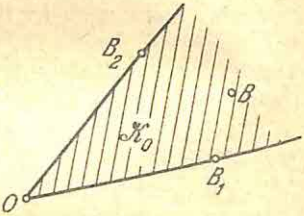


Fig. 37

points of the angle may be written down in the form

$$B = t_1 B_1 + t_2 B_2 \tag{8}$$

where t_1 and t_2 are arbitrary nonnegative numbers; and to find the points B_1 and B_2 is not difficult at all, if it is remembered that each of them (i) belongs to $\mathscr{K}_0$, i.e. satisfies the system (2), and (ii) lies on the boundary of $\mathscr{K}_0$, i.e. satisfies one of the equations (3). If $\mathscr{K}_0$ is a ray, then instead of (8) we have

$$B = tB_1 \tag{9}$$

where B_1 is any point of the ray (different from the origin) and t is an arbitrary nonnegative number.

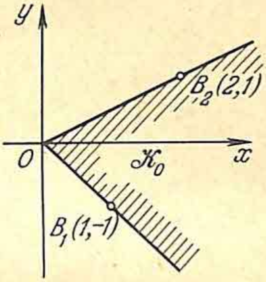


Fig. 38

Example 2. Find the region $\mathscr{K}_0$ of the solutions of the system

$$\left.\begin{aligned} x + y &\geqslant 0 \\ x - 2y &\geqslant 0 \\ 2x - y &\geqslant 0 \end{aligned}\right\} \tag{10}$$

as well as the region $\mathscr{K}$ of the solutions of the system in Example 1.

Solution. The system (10) is normal; for the unique solution of the corresponding homogeneous system of equations

$$\left.\begin{aligned} x + y &= 0 \\ x - 2y &= 0 \\ 2x - y &= 0 \end{aligned}\right\} \tag{11}$$

is (0, 0).

Choose a point satisfying the first equation of (11) (but different from (0, 0)), the point $C(-1, 1)$, for example. We make sure by a simple check that the point C does not satisfy all the inequalities (10), therefore, neither the point itself nor any point of the ray OC (different from the origin O) belongs to $\mathscr{K}_0$. On considering the point $-C$ (i.e. the point $(1, -1)$) we find that it belongs to $\mathscr{K}_0$. So $B_1 = (1, -1)$. The second equation is satisfied by the point (2, 1); it is also a solution of the system (10), so that $B_2 = (2, 1)$.

The region $\mathscr{K}_0$ consists of the points (Fig. 38)

$$t_1B_1 + t_2B_2 = t_1(1, -1) + t_2(2, 1) = (t_1 + 2t_2, -t_1 + t_2)$$

where t_1 and t_2 are arbitrary nonnegative numbers.

Turning to the system of inequalities of Example 1 we notice that it is (10) that is the corresponding homogeneous system. From

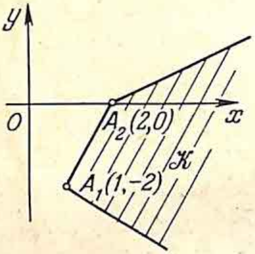


Fig. 39

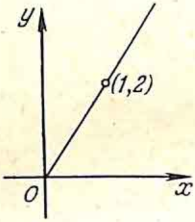


Fig. 40

the theorem proved above we have

$$\mathscr{K} = \langle A_1, A_2 \rangle + \mathscr{K}_0$$

where $A_1(1, -2)$ and $A_2(2, 0)$ are the vertices of the region $\mathscr{K}$. So $\mathscr{K}$ consists of the points (Fig. 39)

$$s(1, -2) + (1 - s)(2, 0) + (t_1 + 2t_2, -t_1 + t_2) =$$
$$= (2 - s + t_1 + 2t_2, -2s - t_1 + t_2)$$

where s is any number in the interval $[0, 1]$ and t_1, t_2 are any nonnegative numbers.

Example 3. Find the feasible region of the system

$$\left.\begin{aligned} 2x - y &\geqslant 0 \\ -4x + 2y &\geqslant 0 \\ x + y &\geqslant 0 \end{aligned}\right\}$$

Proceeding as in Example 2 we find only a single ray:

$$B = t(1, 2) = (t, 2t) \quad (t \geqslant 0)$$

(Fig. 40).

Example 4. Find the feasible region of the system

$$\left.\begin{aligned} 2x - y &\geqslant 0 \\ x + y &\geqslant 0 \\ -3x + y &\geqslant 0 \end{aligned}\right\}$$

In this case none of the equations

$$\begin{aligned} 2x - y &= 0 \\ x + y &= 0 \\ -3x + y &= 0 \end{aligned}$$

has solutions (except (0, 0)) which would satisfy all the given inequalities. The region $\mathscr{K}_0$ consists of a single point (0, 0), the origin of coordinates.

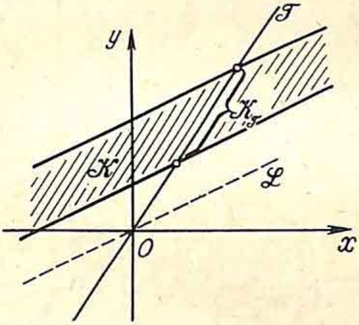


Fig. 41

4°. *The case where the system of inequalities* (1) *is not normal.* This means that the feasible region $\mathscr{L}$ of the homogeneous system of equations (3) contains something else besides the origin of coordinates. Therefore, all the equations (3) determine the same straight line in the plane, and that line is $\mathscr{L}$.

According to Lemma 1 the region $\mathscr{K}$ contains together with each of its points P a straight line $P + \mathscr{L}$ (a line passing through P and parallel to $\mathscr{L}$). Consider a straight line $\mathscr{T}$ not parallel to $\mathscr{L}$. If we know which points of the line $\mathscr{T}$ belong to the region $\mathscr{K}$—denote the set of these points by $\mathscr{K}_{\mathscr{T}}$—then we shall be able to find the region $\mathscr{K}$ itself, for then $\mathscr{K} = \mathscr{K}_{\mathscr{T}} + \mathscr{L}$ (Fig. 41).

The equation of the straight line $\mathscr{L}$ is $a_1x + b_1y = 0$. One of the coefficients a_1 or b_1 of this equation is different from zero; let $b_1 \neq 0$, for example. One can then take as a straight line $\mathscr{T}$ not parallel to $\mathscr{L}$ the y-axis (its equation is $x = 0$). In this case the set $\mathscr{K}_{\mathscr{T}}$—now denote it by $\mathscr{K}_y$—is the part of the y-axis contained by $\mathscr{K}$. To find this set, one should put $x = 0$ in the system (1).

Then we have the system of inequalities

$$\left.\begin{array}{r} b_1y + c_1 \geqslant 0 \\ b_2y + c_2 \geqslant 0 \\ \dots\dots\dots \\ b_my + c_m \geqslant 0 \end{array}\right\} \qquad (12)$$

in one unknown y which is easy to solve*. Notice that the set $\mathscr{K}_y$ may be either an empty set (then $\mathscr{K}$ is also empty) or a point, a segment or a ray (but it cannot be the whole of the y-axis, for otherwise $\mathscr{K}$ would be the whole of the plane, which is impossible). On finding this set we shall know the region $\mathscr{K}$ itself, for

$$\mathscr{K} = \mathscr{K}_y + \mathscr{L} \qquad (13)$$

(if $\mathscr{L}$ is not parallel to the y-axis).

Example 5. Find the feasible region $\mathscr{K}$ of the system

$$\left.\begin{array}{r} x + y - 1 \geqslant 0 \\ -x - y + 2 \geqslant 0 \\ 2x + 2y + 3 \geqslant 0 \end{array}\right\}$$

It is easily seen that the system is not normal and that $\mathscr{L}$ is the straight line

$$x + y = 0$$

(not parallel to the y-axis). Setting $x = 0$, we have the system

$$\left.\begin{array}{r} y - 1 \geqslant 0 \\ -y + 2 \geqslant 0 \\ 2y + 3 \geqslant 0 \end{array}\right\}$$

from which it is seen that $\mathscr{K}_y$, the intersection of $\mathscr{K}$ with the y-axis, is a segment with the ends $C_1(0, 1)$ and $C_2(0, 2)$. So $\mathscr{K}$ is a set of points of the form (Fig. 42)

$$(0, y) + (x, -x) = (x, y - x)$$

where x is arbitrary and y is any number in the interval from 1 to 2.

---

* Notice that the system (12) (viewed as a system of inequalities in one unknown) is now normal. Indeed, if otherwise, the corresponding homogeneous system would have a nonzero solution y^*; but then the system (3) would have the solution $(0, y^*)$ not belonging to $\mathscr{L}$.

In conclusion we shall briefly discuss a theorem which follows from the results obtained above. In the two-dimensional case we are considering (i.e. where everything takes place in the plane) this theorem is not particularly striking and it would be right to regard it as a starting point of the extension to the "n-dimensional" case to be studied in Section 7.

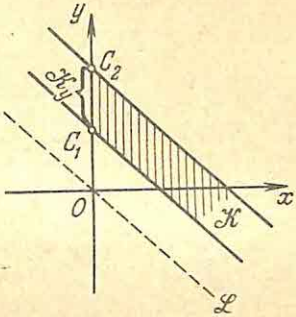


Fig. 42

Theorem. *Any (nonempty) convex polygonal region $\mathcal{K}$ in the plane can be represented as the sum*

$$\langle A_1, A_2, \ldots, A_p\rangle + (B_1, B_2, \ldots, B_q) \qquad (14)$$

The augend of the sum is the convex hull of a system of points A_1, A_2, ..., A_p, and the addend is a set of points of the form $t_1B_1 + t_2B_2 + \ldots + t_qB_q$, where $t_1, t_2, \ldots, t_q$ are arbitrary nonnegative numbers.

The theorem can be proved in a few words. Consider a system of inequalities giving $\mathcal{K}$. If the system is normal, then equation (7) holds; bearing in mind that in this equation $\mathcal{K}_0$ is one of the sets of the form (B_1, B_2), (B_1) or (O) (the origin of coordinates), we find that in the case where the system is normal our statement is valid. If the system is not normal, then equation (13) holds from which it also follows that $\mathcal{K}$ is representable in the required form (why?).

Notice that if all points $A_1, A_2, \ldots, A_p$ coincide with the origin O, then the set $\langle A_1, A_2, \ldots, A_p\rangle$ also coincides with O; then only the addend is left of the sum (14). When the points $B_1, B_2, \ldots, B_q$ coincide with O, the set $(B_1, B_2, \ldots, B_q)$ also coincides with O and only the augend is left of the sum (14).

The converse theorem also holds, though with some reservation.

Theorem. *Any set of the form*

$$\langle A_1, A_2, \dots, A_p\rangle + (B_1, B_2, \dots, B_q)$$

in the plane is either the whole plane or some convex polygonal region in it.

The proof is fairly obvious. The addend, i.e. the region $\mathcal{K}_0 = (B_1, B_2, \dots, B_q)$, is either the whole plane, or a half-plane, or an angle (smaller than 180°), or a ray, or a point (the origin of coordinates),

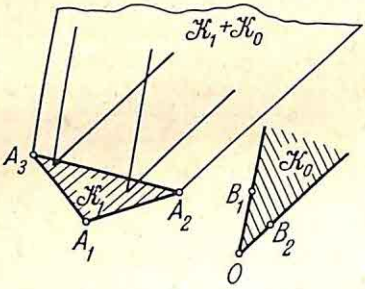


Fig. 43

whereas the summand $\mathcal{K}_1 = \langle A_1, A_2, \dots, A_p\rangle$ represents a convex polygon. The set $\mathcal{K}_1 + \mathcal{K}_0$ can be obtained by translating $\mathcal{K}_0$ to the segments OK_1 (where K_1 is any point of $\mathcal{K}_1$) and taking the union of the sets obtained (Fig. 43). It is easily seen that this produces either the whole plane (this is the case if $\mathcal{K}_0$ is the whole plane) or some convex polygonal region in it.

6. The Feasible Region of a System in Three Unknowns

After the detailed analysis given in the preceding section we are now in a position to minimize the required theory when considering systems in three unknowns.

Along with the original system

$$\left.\begin{array}{c} a_1x + b_1y + c_1z + d_1 \geqslant 0 \\ \dots\dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz + d_m \geqslant 0 \end{array}\right\} \qquad (1)$$

as in Section 5, we again consider two more systems:

$$\left.\begin{array}{l} a_1x + b_1y + c_1z \geqslant 0 \\ \cdots\cdots\cdots\cdots \\ a_mx + b_my + c_mz \geqslant 0 \end{array}\right\} \qquad (2)$$

and

$$\left.\begin{array}{l} a_1x + b_1y + c_1z = 0 \\ \cdots\cdots\cdots\cdots \\ a_mx + b_my + c_mz = 0 \end{array}\right\} \qquad (3)$$

We shall again denote the feasible region of the system (1) by $\mathscr{K}$, that of the system (2) by $\mathscr{K}_0$ and that of the system (3) by $\mathscr{L}$. To use the previously introduced terminology, $\mathscr{K}$ is some convex polyhedral region in space and $\mathscr{K}_0$ is a convex polyhedral cone. As already noted, Lemmas 1 and 2 of Section 5 remain valid here too.

1°. *The case where the system of inequalities* (1) *is normal.* Here the region $\mathscr{K}$ does not contain any straight lines and hence has at least one vertex. Indeed, if $\mathscr{K}$ lies in the plane (and, as noted in Section 2, such a case is possible), then $\mathscr{K}$ is a convex polygonal region in the plane which contains no straight lines and, as explained in Subsection 2° of Section 5, must, therefore, have vertices. If the region $\mathscr{K}$ does not lie in the plane, however, consider its boundary. This consists of faces each of which, being a convex polygonal region containing no straight lines, must have vertices; and it is easily seen that a vertex of any face is simultaneously a vertex of the region $\mathscr{K}$.

Converging in the vertex A of the region $\mathscr{K}$ are at least three boundary planes for which the point A is the *only* common point. Indeed, if this were not the case, all the boundary planes passing throught A would either coincide or have a common straight line. But in that case a sufficiently small segment passing through A and lying in the common boundary plane or on the common boundary line would belong to $\mathscr{K}$, which is contrary to the definition of a vertex.

This compels us to make some obvious changes in the method of finding vertices described in Subsection 2° of Section 5. That is, now we should say that a *regular subsystem* is a subsystem of *three* equations of the system

$$\left.\begin{array}{l} a_1x + b_1y + c_1z + d_1 = 0 \\ \cdots\cdots\cdots\cdots \\ a_mx + b_my + c_mz + d_m = 0 \end{array}\right\} \qquad (4)$$

rather than two, provided the solution (x, y, z) of the subsystem is unique. When a regular system is understood like this, the method of finding vertices remains exactly the same as before, namely:

To find all the vertices of the region $\mathscr{K}$, one should find the solutions of all the regular subsystems of the system (4) *and pick out those which satisfy the original system* (1).

The theorem of Subsection 2° of Section 5 also remains valid; the changes to be made in the proof of the theorem are obvious. The remark that a normal system has no solutions if the region $\mathscr{K}$ has no vertices also remains valid.

Example 1. Find the vertices of the region $\mathscr{K}$ given by the system of inequalities

$$\left.\begin{aligned} 2x + y + z - 1 &\geqslant 0 \\ x + 2y + z - 1 &\geqslant 0 \\ x + y + 2z - 1 &\geqslant 0 \\ x + y + z - 1 &\geqslant 0 \end{aligned}\right\} \tag{5}$$

Here the corresponding homogeneous system of equations is of the form

$$\left.\begin{aligned} 2x + y + z &= 0 \\ x + 2y + z &= 0 \\ x + y + 2z &= 0 \\ x + y + z &= 0 \end{aligned}\right\}$$

Solving it we see that the unique solution is (0, 0, 0); so the system (5) is normal.

To find the vertices we shall have to consider all the subsystems of three equations of the system (4):

$$\left.\begin{aligned} 2x + y + z - 1 &= 0 \\ x + 2y + z - 1 &= 0 \\ x + y + 2z - 1 &= 0 \end{aligned}\right\} \qquad \left.\begin{aligned} 2x + y + z - 1 &= 0 \\ x + 2y + z - 1 &= 0 \\ x + y + z - 1 &= 0 \end{aligned}\right\}$$

$$\left.\begin{aligned} 2x + y + z - 1 &= 0 \\ x + y + 2z - 1 &= 0 \\ x + y + z - 1 &= 0 \end{aligned}\right\} \qquad \left.\begin{aligned} x + 2y + z - 1 &= 0 \\ x + y + 2z - 1 &= 0 \\ x + y + z - 1 &= 0 \end{aligned}\right\}$$

On making the required calculations we find that all the subsystems are regular and that their solutions are the points

$$(1/4, 1/4, 1/4), \quad (0, 0, 1), \quad (0, 1, 0), \quad (1, 0, 0)$$

of which the first does not satisfy the system (5) and the remaining three do. Therefore, the vertices of the region $\mathscr{K}$ are the following:

$$A_1(1, 0, 0), \quad A_2(0, 1, 0), \quad A_3(0, 0, 1)$$

2°. *The normal homogeneous system of inequalities* (2). Each of the inequalities (2) determines a half-space whose boundary plane passes through the origin of coordinates.

Here a single point, the origin of coordinates, is the intersection of the boundary planes (the system (2) is normal!). In other words, the set $\mathscr{K}_0$, the feasible region of the system (2), is a convex polyhedral cone with a *single* vertex. It follows from the enumeration of convex polyhedral cones given in Section 4 that in our case $\mathscr{K}_0$ is either an infinite convex pyramid, a flat angle, a ray or, finally, a single point (the origin of coordinates). Omitting the last case for the time being, we have in all the remaining ones

$$\mathscr{K}_0 = (B_1, B_2, ..., B_q)$$

where B_1, B_2, ..., B_q are some points chosen one at a time on each edge of the cone $\mathscr{K}_0$ (see Theorem 2 of Section 4). One can find such points from the following considerations. Each of them (i) belongs to $\mathscr{K}_0$, i.e. satisfies the system (2), and (ii) belongs to the line of intersection of two different faces, i.e. satisfies disproportionate* equations of the system (3).

If it is found that the only point satisfying conditions (i) and (ii) is (0, 0, 0), then the region $\mathscr{K}_0$ coincides with the origin of coordinates.

Example 2. Find the region $\mathscr{K}_0$ of the solutions of the system

$$\left.\begin{aligned} 2x + y + z &\geqslant 0 \\ x + 2y + z &\geqslant 0 \\ x + y + 2z &\geqslant 0 \\ x + y + z &\geqslant 0 \end{aligned}\right\} \qquad (6)$$

and, further, the region $\mathscr{K}$ of the solutions of the system in Example 1.

Notice first of all that the system (6) is connected with the system of inequalities (5) of Example 1; namely (6) is a homogeneous system corresponding to (5). Hence the system (6) is normal.

---

* We call the equations $ax + by + cz = 0$ and $a'x + b'y + c'z = 0$ "disproportionate" if at least one of the equalities $a/a' = b/b' = c/c'$ fails to hold; in this case the corresponding planes intersect along a straight line.

Here the system of two disproportionate equations can be set up in six different ways:

$$\left.\begin{aligned}x+2y+z&=0\\x+y+2z&=0\end{aligned}\right\}\quad\left.\begin{aligned}2x+y+z&=0\\x+y+2z&=0\end{aligned}\right\}\quad\left.\begin{aligned}2x+y+z&=0\\x+y+z&=0\end{aligned}\right\}$$

$$\left.\begin{aligned}2x+y+z&=0\\x+2y+z&=0\end{aligned}\right\}\quad\left.\begin{aligned}x+2y+z&=0\\x+y+z&=0\end{aligned}\right\}\quad\left.\begin{aligned}x+y+2z&=0\\x+y+z&=0\end{aligned}\right\}$$

For each of these six systems, choose two nonzero solutions: (x, y, z) and $(-x, -y, -z)$. For example, one can take $(3, -1, -1)$ and $(-3, 1, 1)$ for the first system; inequalities (6) are satisfied only by the first of these solutions. Hence we have the point $B_1 = (3, -1, -1)$. Proceeding in similar fashion with the remaining five systems we find the points $B_2 = (-1, 3, -1)$ and $B_3 = (-1, -1, 3)$. So the region $\mathscr{K}_0$ consists of points of the form

$$t_1B_1 + t_2B_2 + t_3B_3 =$$
$$= (3t_1 - t_2 - t_3, -t_1 + 3t_2 - t_3, -t_1 - t_2 + 3t_3)$$

where t_1, t_2, t_3 are arbitrary nonnegative numbers.

We now turn to the system of inequalities (5) of Example 1. As already noted, it is (6) that is the corresponding homogeneous system. Therefore, the region $\mathscr{K}$ is of the form

$$\langle A_1, A_2, A_3\rangle + \mathscr{K}_0$$

and consists of the points

$$\begin{aligned}&s_1A_1 + s_2A_2 + s_3A_3 + t_1B_1 + t_2B_2 + t_3B_3 =\\&\qquad = s_1(1, 0, 0) + s_2(0, 1, 0) + s_3(0, 0, 1) +\\&+ t_1(3, -1, -1) + t_2(-1, 3, -1) + t_3(-1, -1, 3) =\\&= (s_1 + 3t_1 - t_2 - t_3, s_2 - t_1 + 3t_2 - t_3, s_3 - t_1 - t_2 + 3t_3)\end{aligned}$$

where the numbers t_1, t_2, t_3 are arbitrary nonnegative ones and s_1, s_2, s_3 are nonnegative and sum to unity.

3°. *The case where the system of inequalities* (1) *is not normal.* This means that the feasible region $\mathscr{L}$ of the homogeneous system of equations (3) contains points different from the origin of coordinates. Since $\mathscr{L}$ is the intersection of planes, two cases are possible:

1. $\mathscr{L}$ is a straight line. According to Lemma 1 the region $\mathscr{K}$ together with each of its points P contains a straight line $P + \mathscr{L}$. Consider some plane $\mathscr{T}$ which is not parallel to $\mathscr{L}$. If we know which points of the plane $\mathscr{T}$ belong to the region $\mathscr{K}$—denote

the set of these points by $\mathscr{K}_{\mathscr{T}}$—we shall be able to find the region $\mathscr{K}$ itself, for then $\mathscr{K} = \mathscr{K}_{\mathscr{T}} + \mathscr{L}$.

But, whatever the straight line $\mathscr{L}$ may be, one can always choose one of the coordinate planes, xOy, xOz or yOz, as the plane $\mathscr{T}$ not parallel to it. Assume, for example, that $\mathscr{L}$ is not parallel to the yOz plane. Take this plane to be $\mathscr{T}$. Then the set $\mathscr{K}_{\mathscr{T}}$ – denote it now by $\mathscr{K}_{y,z}$—is the part of the yOz plane contained in $\mathscr{K}$ (Fig. 44). To find this set, one should put $x = 0$ in the system (1). We then have the system of inequalities*

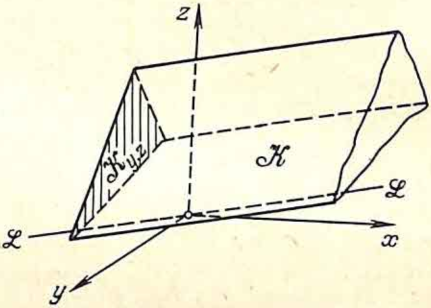


Fig. 44

$$\left.\begin{array}{l} b_1 y + c_1 z + d_1 \geqslant 0 \\ \dots\dots\dots\dots \\ b_m y + c_m z + d_m \geqslant 0 \end{array}\right\} \qquad (7)$$

which can be solved using the method of Section 5.

On finding the set $\mathscr{K}_{y,z}$ we shall be able to write down

$$\mathscr{K} = \mathscr{K}_{y,z} + \mathscr{L} \qquad (8)$$

(if the straight line $\mathscr{L}$ is not parallel to the yOz plane), which is the complete description of the region $\mathscr{K}$.

Remark. If it is found that the set $\mathscr{K}_{y,z}$ is empty, so is $\mathscr{K}$. This means that the system (1) is incompatible.

Example 3. Find the region $\mathscr{K}$ of the solutions of the system

$$\left.\begin{array}{l} -2x + y + z - 1 \geqslant 0 \\ -3x - y + 4z - 1 \geqslant 0 \\ -x - 2y + 3z \geqslant 0 \end{array}\right\} \qquad (9)$$

---

* It is easily seen that the system (7) is now normal (see a similar footnote on page 42).

Consider the corresponding homogeneous system of equations

$$\left.\begin{aligned} -2x + y + z &= 0 \\ -3x - y + 4z &= 0 \\ -x - 2y + 3z &= 0 \end{aligned}\right\} \qquad (10)$$

Solving it we find that the third equation is a consequence of the first two, so that the system is reduced to the first two equations. The set of its solutions $\mathscr{L}$ is the straight line along which the planes

$$-2x + y + z = 0$$

and

$$-3x - y + 4z = 0$$

intersect.

Choose some point B on the straight line $\mathscr{L}$ different from the origin of coordinates. To do this, it is enough to find some three numbers x, y, z (not vanishing all together) which satisfy the first two equations of the system (10). Take 1, 1, 1, for example. So $\mathscr{L}$ is the straight line OB, where $B = (1,\ 1,\ 1)$.

It is easy to see that the straight line $\mathscr{L}$ is not parallel to, say, the yOz coordinate plane. Setting $x = 0$ in the system (9), we have the system

$$\left.\begin{aligned} y + z - 1 &\geqslant 0 \\ -y + 4z - 1 &\geqslant 0 \\ -2y + 3z \qquad &\geqslant 0 \end{aligned}\right\}$$

in two unknowns y and z, which is normal. One can find its feasible region $\mathscr{K}_{y,z}$ using the method of Section 5. On making the required calculations we find that $\mathscr{K}_{y,z}$ is a set consisting of a single point A(3/5, 2/5) (in the yOz plane). Hence the sought region $\mathscr{K}$ consists of all the points of the form

$$A + tB = \left(0,\ \tfrac{3}{5},\ \tfrac{2}{5}\right) + t(1,\ 1,\ 1) = \left(t,\ \tfrac{3}{5} + t,\ \tfrac{2}{5} + t\right)$$

where t is any nonnegative number (the region $\mathscr{K}$ is a straight line parallel to $\mathscr{L}$).

2. $\mathscr{L}$ is a plane. Then take as a secant set $\mathscr{T}$ some straight line not parallel to this plane; in particular, one can take one of the coordinate axes. Assume, for example, that the z-axis is not parallel to $\mathscr{L}$; take it to be $\mathscr{T}$. To find the set $\mathscr{K}_z$, the part of the z-axis contained in $\mathscr{K}$, one should put $x = 0$, $y = 0$ in the system (1).

We then have the system of inequalities

$$\left.\begin{array}{r} c_1 z + d_1 \geqslant 0 \\ \cdots\cdots\cdots \\ c_m z + d_m \geqslant 0 \end{array}\right\} \qquad (11)$$

which is solved without much difficulty*. On finding the set $\mathscr{K}_z$, we shall be able to write down (Fig. 45)

$$\mathscr{K} = \mathscr{K}_z + \mathscr{L} \qquad (12)$$

(if the plane $\mathscr{L}$ is not parallel to the z-axis), which is the complete description of $\mathscr{K}$.

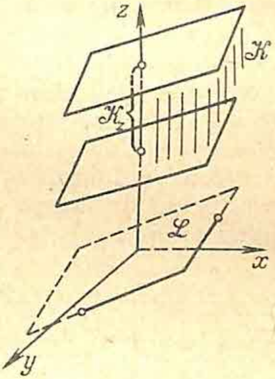


Fig. 45

Remark. If the set $\mathscr{K}_z$ is found to be empty, so is $\mathscr{K}$. The system (1) is then incompatible.

Example 4. Find the region $\mathscr{K}$ of the solutions of the system

$$\left.\begin{array}{r} x - y + z + 1 \geqslant 0 \\ -x + y - z + 2 \geqslant 0 \end{array}\right\} \qquad (13)$$

Here the corresponding homogeneous system of equations is of the form

$$\left.\begin{array}{r} x - y + z = 0 \\ -x + y - z = 0 \end{array}\right\} \qquad (14)$$

---

* The system (11) is normal.

The second equation is a consequence of the first; therefore, the feasible region of the system (14) is the plane $\mathscr{L}$ given by the equation

$$x - y + z = 0$$

It is easy to see that this plane intersects the z-axis in a single point and hence is not parallel to the z-axis. Find the set $\mathscr{K}_z$.

Setting $x = 0$, $y = 0$ in the system (13), we have the system

$$\left.\begin{aligned} z + 1 \geqslant 0 \\ -z + 2 \geqslant 0 \end{aligned}\right\}$$

from which it follows that

$$-1 \leqslant z \leqslant 2 \tag{15}$$

So $\mathscr{K}$ is the set $\mathscr{K}_z + \mathscr{L}$ consisting of the points of the form $(0, 0, z) + (x, y, -x + y) = (x, y, z - x + y)$ where x and y are arbitrary and z satisfies the inequality (15).

We conclude this section by formulating two theorems which generalize the last two theorems of Section 5 to the three-dimensional case. The only change we should make for this purpose in the formulations of the above-mentioned theorems of Section 5 consists in substituting the word "space" for "plane".

Theorem. *Any (nonempty) convex polyhedral region in space can be represented as the sum*

$$\langle A_1, A_2, \ldots, A_p\rangle + (B_1, B_2, \ldots, B_q)$$

Theorem. *Any set of the form*

$$\langle A_1, A_2, \ldots, A_p\rangle + (B_1, B_2, \ldots, B_q)$$

in space is either the whole of space or some convex polyhedral region in it.

The proofs of both theorems follow almost word for word those of the corresponding theorems in the two-dimensional case. Their elaboration is left to the reader.

7. Systems of Linear Inequalities in Any Number of Unknowns

In the foregoing sections we concentrated upon systems of inequalities *in two or three unknowns*. This was dictated mainly by two circumstances. Firstly, these systems are simple to investigate and allow one to keep entirely within the framework of "school" mathematics. Secondly, (which is more important in the present case) solutions of such systems have visual geometrical meaning (as points in the plane or in space). However, more common in applications (in linear programming problems, for example) are systems of inequalities involving $n > 3$ unknowns. To pass them by in silence would greatly water down the presentation of the question. Therefore, we shall try to at least outline the situation for any $n > 3$.

To visualize linear systems of inequalities in n unknowns we must turn to the so-called *n-dimensional space.*

We shall begin with the definition of relevant concepts, confining ourselves only to the most essential.

A point of n-dimensional space is defined by an ordered set of n numbers

$$x_1, x_2, \ldots, x_n$$

called the *coordinates* of the point. Such a definition is motivated by the fact, fundamental to analytic geometry, that a point in the plane is described by a pair of numbers and that in space by a triple of numbers. In what follows instead of saying "the point M has the coordinates $x_1, x_2, \ldots, x_n$" we shall write $M = (x_1, x_2, \ldots, x_n)$ or simply $M(x_1, x_2, \ldots, x_n)$. The point $(0, 0, \ldots, 0)$ is called the *origin of coordinates* or just the *origin.*

We shall first of all point out what is meant by a "segment" in n-dimensional space. According to Section 1, in ordinary space a segment M_1M_2 can be described as a set of all points of the form

$$s_1M_1 + s_2M_2$$

where s_1, s_2 are any two nonnegative numbers whose sum is one. Going from three-dimensional space to n-dimensional space we adopt this description *as the definition of a segment*. More strictly, let

$$M'(x_1', x_2', \ldots, x_n') \text{ and } M''(x_1'', x_2'', \ldots, x_n'')$$

be two arbitrary points of n-dimensional space. Then we say that a *segment $M'M''$ is a set of all points of the form*

$$\begin{aligned} &s'M' + s''M'' = \\ &\quad = (s'x_1' + s''x_1'', s'x_2' + s''x_2'', \ldots, s'x_n' + s''x_n'') \end{aligned} \tag{1}$$

where s', s'' are any two nonnegative numbers with a sum of one. When $s' = 1, s'' = 0$ we have the point M', when $s' = 0, s'' = 1$ we have the point M''. These are the *ends* of the segment $M'M''$. The remaining points (they are obtained when $s' > 0, s'' > 0$) are called the *interior points* of the segment.

Of other concepts pertaining to n-dimensional space we shall need the concept of *hyperplane*. This is a generalization of the concept of plane in ordinary three-dimensional space. Here the prefix "hyper" has quite definite meaning. The fact is that in n-dimensional space various types of "planes" are possible, viz. one-dimensional "planes" (they are called "straight lines"), two-dimensional "planes", etc., and finally $(n-1)$-dimensional "planes": it is the last that are called "hyperplanes".

Definition. *In n-dimensional space, a set of points* $M(x_1, x_2, \dots, x_n)$ *whose coordinates satisfy the equation of the first degree*

$$a_1x_1 + a_2x_2 + \dots + a_nx_n + b = 0 \qquad (2)$$

where at least one of the numbers $a_1, a_2, \dots, a_n$ (coefficients of the unknowns) *is different from zero, is called a hyperplane*. For $n = 3$ the equation (2) assumes the form $a_1x_1 + a_2x_2 + a_3x_3 + b = 0$, which is none other than the equation of a plane in ordinary space (where the coordinates are denoted by x_1, x_2, x_3 instead of x, y, z as usual).

With respect to the hyperplane (2) the whole of n-dimensional space is divided into two parts: the region where the inequality

$$a_1x_1 + a_2x_2 + \dots + a_nx_n + b \geqslant 0 \qquad (3)$$

holds and the region where we have

$$a_1x_1 + a_2x_2 + \dots + a_nx_n + b \leqslant 0 \qquad (4)$$

These regions are called *half-spaces*. Thus every hyperplane divides the whole of space into two half-spaces for which it is a common part.

The concept of a convex solid is also generalized to the n-dimensional case. A set of points in n-dimensional space is said to be *convex* if together with any two of its points M' and M'' it contains the entire segment $M'M''$ as well.

It is easily shown that *any half-space is a convex set*. Indeed, let the points $M'(x'_1, x'_2, \dots, x'_n)$ and $M''(x''_1, x''_2, \dots, x''_n)$ belong to the half-space (3). We prove that any point M of the segment $M'M''$ also belongs to that half-space.

The coordinates of a point M are written in the form (1) or, equivalently, as

$$\begin{aligned} x_1 &= sx'_1 + (1-s)x''_1 \\ x_2 &= sx'_2 + (1-s)x''_2 \\ &\dots\dots\dots\dots \\ x_n &= sx'_n + (1-s)x''_n \end{aligned} \qquad (0 \leqslant s \leqslant 1)$$

Substituting into the left-hand side of (3) we have

$$\begin{aligned} a_1(sx_1' + (1-s)x_1'') + a_2(sx_2' + (1-s)x_2'') + \ldots \\ \ldots + a_n(sx_n' + (1-s)x_n'') + b = \\ = s(a_1x_1' + a_2x_2' + \ldots + a_nx_n') + \\ + (1-s)(a_1x_1'' + a_2x_2'' + \ldots + a_nx_n'') + sb + (1-s)b \end{aligned}$$

(we have replaced the number b by the sum $sb + (1-s)b$), which is equal to

$$s[a_1x_1' + \ldots + a_nx_n' + b] + (1-s)[a_1x_1'' + \ldots + a_nx_n'' + b]$$

Each of the two sums in square brackets is nonnegative since both points M' and M'' belong to the half-space (3). Therefore the entire expression is also nonnegative (for $s \geqslant 0$ and $(1-s) \geqslant 0$). This proves that the point M belongs to the half-space (3), i.e. that this half-space is convex.

From the above it is easy to understand what geometric terminology should correspond to a system of linear inequalities in n unknowns. Given the system

$$\begin{aligned} a_1x_1 + a_2x_2 + \ldots + a_nx_n + a \geqslant 0 \\ b_1x_1 + b_2x_2 + \ldots + b_nx_n + b \geqslant 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ c_1x_1 + c_2x_2 + \ldots + c_nx_n + c \geqslant 0 \end{aligned} \qquad (5)$$

each of the inequalities determines a certain half-space while all the inequalities together determine a certain region $\mathcal{K}$ in n-dimensional space which is the intersection of a finite number of half-spaces. The region $\mathcal{K}$ is convex, since so is any of the half-spaces that form it.

In n-dimensional space, by analogy with the three-dimensional case, *a region which is the intersection of a finite number of half-spaces is called a convex polyhedral region or, if the intersection is a bounded set, simply a convex polyhedron.* Here the words "bounded set" should be understood in the sense that the coordinates of all the points of a region under consideration do not exceed in absolute value a certain constant c:

$|x_1| \leqslant c, \ldots, |x_n| \leqslant c$ for all points of a given region.

Thus, *in n-dimensional space, a set of points whose coordinates satisfy the system* (5) *is the convex polyhedral region* $\mathcal{K}$ *resulting from intersection of all half-spaces corresponding to the inequalities of the given system.*

Note again that if the region is bounded, it is called a convex polyhedron.

The methods of actual description of the region $\mathscr{K}$ which have been considered for systems in two unknowns in Section 5 and for systems in three unknowns in Section 6 can, with appropriate changes, be extended to the case of n unknowns. We shall not go into this, however, for it would require rather much space. Besides, with a large number of unknowns these methods are not efficient, requiring as they do, too cumbersome calculations.

It is remarkable that the general theorems on the structure of convex polyhedral sets in three-dimensional space remain fully valid in n-dimensional space, although requiring more elaborate proofs. We shall restrict ourselves only to the formulations of the theorems and to the essential explanations.

Theorem 1. *The convex hull of any finite system of the points $A_1, A_2, \ldots, A_p$ is a convex polyhedron.*

For better intelligibility we shall emphasize that here a relation is established between two differently defined types of sets, viz. between the convex hull of a system of points $A_1, A_2, \ldots, A_p$ which is designated as $\langle A_1, A_2, \ldots, A_p \rangle$ and defined as a set of all points of the form

$$s_1A_1 + s_2A_2 + \ldots + s_pA_p$$

where $s_1, s_2, \ldots, s_p$ are any nonnegative numbers whose sum is one, and convex polyhedra, i.e. bounded regions resulting from intersection of a finite number of half-spaces.

In two- and three-dimensional spaces the validity of Theorem 1 is obvious (at least from the visualization of the convex hull), while in the multidimensional case it is not obvious at all and requires to be proved.

Theorem 1′ (the converse of Theorem 1). *Any convex polyhedron coincides with the convex hull of some finite system of points.*

In fact we can state even more: *a convex polyhedron coincides with the convex hull of its vertices.* The definition of a vertex is the same as in the two-dimensional case (a vertex is such a point of a polyhedron which is not an interior point for any of the segments belonging wholly to the polyhedron). It can be shown that the number of vertices is always finite.

Theorem 2. *Any set of the form $(B_1, B_2, \ldots, B_q)$ either coincides with the whole of space or is some convex polyhedral cone with vertex at origin.*

We remind that the symbol $(B_1, B_2, \ldots, B_q)$ designates a set of all points which can be represented in the form

$$t_1B_1 + t_2B_2 + \dots + t_qB_q$$

where $t_1, t_2, \dots, t_q$ are any nonnegative numbers. A convex polyhedral cone is defined as the intersection of a finite number of half-spaces whose boundary hyperplanes have a common point (the vertex of the cone). The validity of Theorem 2 in three-dimensional space was established in Section 4 (Theorem 1 of Section 4).

Theorem 2′. *Any convex polyhedral cone with vertex at origin can be represented as* $(B_1, B_2, \dots, B_q)$.

The validity of the theorem for the three-dimensional case was proved in Section 4 (Theorem 2 of Section 4).

Theorem 3. *Any convex polyhedral region can be represented as the sum*

$$\langle A_1, A_2, \dots, A_p\rangle + (B_1, B_2, \dots, B_q)$$

Theorem 3′. *Any sum of the indicated form is either the whole of space or some convex polyhedral region in it.*

8. The Solution of a System of Linear Inequalities by Successive Reduction of the Number of Unknowns

From his elementary algebra course reader is familiar with the method for solving systems of linear *equations* by successive reduction of the number of unknowns. For systems involving three unknowns x, y, z the essence of the method can be described as follows.

From each equation of a given system, one finds the unknown z in terms of the unknowns x and y. The obtained expressions (containing only x and y) are then equated to one another. This leads to a new system, a system in two unknowns x and y. For example, the following was the original system

$$\left.\begin{aligned} 2x - 3y + z &= -1 \\ x - y + 2z &= 5 \\ 4y - z &= 5 \end{aligned}\right\} \qquad (1)$$

then, on solving each of the equations for z, we have

$$\left.\begin{aligned} z &= -2x + 3y - 1 \\ z &= -\tfrac{1}{2}x + \tfrac{1}{2}y + \tfrac{5}{2} \\ z &= 4y - 5 \end{aligned}\right\} \qquad (1')$$

after equating the expressions in the right-hand sides (it is enough to equate the first of the expressions to each of the rest in turn) we have the system

$$\left.\begin{aligned} -2x+3y-1 &= -\tfrac{1}{2}x+\tfrac{1}{2}y+\tfrac{5}{2} \\ -2x+3y-1 &= \qquad\quad 4y-5 \end{aligned}\right\} \tag{2}$$

in two unknowns x and y. The system (2) can be handled in the same way. From the system of equations (2) we find

$$\left.\begin{aligned} y &= \quad \tfrac{3}{5}x+\tfrac{7}{5} \\ y &= -2x+4 \end{aligned}\right\} \tag{2'}$$

after which we get the equation

$$\tfrac{3}{5}x+\tfrac{7}{5}=-2x+4$$

involving now one unknown x. On solving the equation we find $x=1$, after which from the system (2') we find $y=2$, and finally from the system (1') we have $z=3$.

It turns out that something of the kind can be done with any system of linear *inequalities*. Solving a system of inequalities in n unknowns $x_1, \ldots, x_{n-1}, x_n$ is thus reduced to solving a system of inequalities in $n-1$ unknowns $x_1, \ldots, x_{n-1}$, then the resulting system can be reduced to a system in $n-2$ unknowns $x_1, \ldots, x_{n-2}$ and so on and on till we come to a system in one unknown x_1, and solving a system in one unknown is quite an elementary task. In this way we get an opportunity of finding any solution (at least in principle) of the original system of inequalities.

So suppose a system of linear inequalities in n unknowns $x_1, x_2, \ldots, x_n$ is given (for reference convenience it will be referred to as "the system (S)" in what follows).

A number of questions arise in discussing the system (S): Is it compatible? If it is, how are all of its solutions to be found? When are systems of inequalities incompatible? Answers to all these questions will be given in this and the next section. Here it is found very helpful to connect to each system of inequalities a new system in which the number of unknowns is less by one than in the original system; this new system is called *concomitant*.

We proceed to describe it.

Consider any of the inequalities of the system (*S*). It is of the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n + a \geqslant 0 \tag{3}$$

If $a_n = 0$, leave the inequality unchanged. If $a_n < 0$, transfer the term a_nx_n to the right-hand side of the inequality and divide both sides by a positive number $-a_n$ to get the inequality of the form

$$b_1x_1 + \dots + b_{n-1}x_{n-1} + b \geqslant x_n$$

In the case of $a_n > 0$, transfer all the summands but a_nx_n to the right-hand side of the inequality and divide both sides by a_n to get the inequality

$$x_n \geqslant c_1x_1 + \dots + c_{n-1}x_{n-1} + c$$

So on multiplying each of the inequalities of the original system by an appropriate positive number we get the *equivalent* system of the form

$$\begin{array}{c} \boxed{\begin{array}{c} P_1 \geqslant x_n \\ P_2 \geqslant x_n \\ \cdot\ \cdot\ \cdot\ \cdot \\ P_p \geqslant x_n \end{array}} \\ \boxed{\begin{array}{c} x_n \geqslant Q_1 \\ x_n \geqslant Q_2 \\ \cdot\ \cdot\ \cdot\ \cdot \\ x_n \geqslant Q_q \end{array}} \\ \boxed{\begin{array}{c} R_1 \geqslant 0 \\ R_2 \geqslant 0 \\ \cdot\ \cdot\ \cdot\ \cdot \\ R_r \geqslant 0 \end{array}} \end{array} \tag{T}$$

where $P_1, \dots, P_p$, $Q_1, \dots, Q_q$, $R_1, \dots, R_r$ are expressions of the form $a_1x_1 + \dots + a_{n-1}x_{n-1} + a$ (not containing x_n)*.

---

*Of course, if the system (*S*) has no inequality in which $a_n < 0$, then the system (*T*) will not have the first block. Similarly, if there are no inequalities with $a_n > 0$, then the system (*T*) will not have the second block. Finally, if (*S*) has no inequalities with $a_n = 0$, then the third block will be missing from (*T*).

In shorthand notation the system (T) can be written down as

$$\left.\begin{array}{r} P_\alpha \geqslant x_n \geqslant Q_\beta \\ R_\gamma \geqslant 0 \end{array}\right\}$$

where α is any of the numbers $1, 2, \ldots, p$, β is any of the numbers $1, 2, \ldots, q$, and γ is any of the numbers $1, 2, \ldots, r$.

Along with the system (T) consider the system

$$\left.\begin{array}{r} P_\alpha \geqslant Q_\beta \\ R_\gamma \geqslant 0 \end{array}\right\} \qquad (S')$$

(where α is any of the numbers $1, 2, \ldots, p$, β is any of the numbers $1, 2, \ldots, q$, and γ is any of the numbers $1, 2, \ldots, r$) in $n-1$ unknowns $x_1, \ldots, x_{n-1}$*. Let us agree to call this system *concomitant* (with respect to the original system (S) or to the equivalent system (T)).

There is a close connection between the solutions of the systems (S) and (S') which is expressed in the following

Theorem. *If the value of the last unknown x_n is discarded from any solution of the system (S), then we get a certain solution of the concomitant system (S').*

Conversely, for any solution of the concomitant system (S') it is possible to find such a value of the unknown x_n on adjoining which we get a solution of the original system (S).

The first statement of the theorem is obvious (if a set of values of the unknowns satisfies the system (S), then it does the system (T) too; but then all the inequalities of the system (S') hold for the same set too). We prove the second statement.

Let

$$x_1 = x_1^0, \ldots, x_{n-1} = x_{n-1}^0$$

be a solution of the system (S'). Substituting into the expressions $P_1, \ldots, P_p$, $Q_1, \ldots, Q_q$, $R_1, \ldots, R_r$, we get some numbers $P_1^0, \ldots, P_p^0$, $Q_1^0, \ldots, Q_q^0$, $R_1^0, \ldots, R_r^0$. The following inequalities must hold for them

$$P_\alpha^0 \geqslant Q_\beta^0 \qquad (4)$$

(α being any of the numbers $1, 2, \ldots, p$, and β any of the numbers 1, 2,..., q) and

---

* If the first or the second block is found to be missing from the system (T), then (S') will consist only of the inequalities $R_\gamma \geqslant 0$. If the third block is missing from (T), then (S') will have only inequalities $P_\alpha \geqslant Q_\beta$.

$$R_{\gamma}^{0} \geqslant 0 \tag{5}$$

(γ being any of the numbers 1, 2,..., r).

The first group of the above inequalities (i. e. (4)) shows that each of the numbers $Q_1^0, \ldots, Q_q^0$ is not greater than any of the numbers $P_1^0, \ldots, P_p^0$. But in such a case there must be a number x_n^0 which lies between all the numbers $Q_1^0, \ldots, Q_q^0$ and all the numbers $P_1^0, \ldots, P_p^0$:

$$P_{\alpha}^{0} \geqslant x_n^0 \geqslant Q_{\beta}^{0}$$

where α is any of the numbers 1, 2,..., p, and β is any of the numbers 1, 2,..., q (Fig. 46). These inequalities together with the inequalities (5) mean that the set of the values of the unknowns

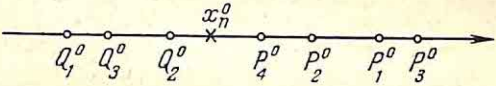


Fig. 46

$$x_1 = x_1^0, \ldots, x_{n-1} = x_{n-1}^0,\ x_n = x_n^0$$

is a solution of the system (T) and hence of (S). Thus the theorem is proved.

The following two additions to the theorem will play an important role below.

1. *A system* (S) *of linear inequalities is compatible if and only if so is the concomitant system* (S'). This is a direct consequence of the theorem proved.

2. *All the solutions of the original system* (S) *can be obtained in the following way. One must adjoin to each solution* $x_1^0, \ldots, x_{n-1}^0$ *of the concomitant system* (S') *any of the numbers* x_n^0 *lying between all the numbers* $Q_1^0, \ldots, Q_q^0$ *and all the numbers* $P_1^0, \ldots, P_p^0$. As a matter of fact this proposition was proved during the proof of the theorem.

We shall say a few words about the geometric meaning of the theorem proved above. Suppose that (S) is a system of inequalities in *three* unknowns x, y, z. The concomitant system (S') is a system in two unknowns x, y. Denote by $\mathscr{K}$ (S) the feasible region of the system (S) (it is a certain set of points in space) and by $\mathscr{K}$ (S') the feasible region of the system (S') (a set of points in the plane). In geometric terms the theorem proved has the following meaning.

The region $\mathscr{K}$ (S') *is the projection of the region* $\mathscr{K}$ (S) *on the* xOy *coordinate plane.*

So for an arbitrary system (S) of linear inequalities involving unknowns $x_1, x_2, \ldots, x_n$ we have constructed a new, concomitant system (S') involving the unknowns $x_1, x_2, \ldots, x_{n-1}$, and for the system (S') one can in turn construct the concomitant system (S'') (involving the unknowns $x_1, x_2, \ldots, x_{n-2}$), for the latter system one can construct the concomitant system (S''') and so on.

Continuing the process we shall, after a number of steps, come to the system $(S^{(n-1)})$ consisting of inequalities in *one* unknown x_1. It follows from Proposition 1 stated above that the system (S) is compatible if and only if so is the system $(S^{(n-1)})$, and answering the question concerning the compatibility or otherwise of a system in one unknown presents no difficulty. Thus it becomes possible for us to learn by quite simple calculations whether the system (S) is compatible or not.

Suppose that the system is compatible. Then the problem arises of solving the system or, speaking at greater length, of *indicating all of its solutions* (all sets of values of the unknowns which satisfy the inequalities of the given system). We shall adopt the point of view (though it may at first appear to the reader strange) that the *system* (S) *is solved if the systems* $(S'), (S''), \ldots, (S^{(n-1)})$ *are constructed.* We shall presently see what explains such a point of view but first we shall introduce a

Definition. A set of values of the first k unknowns

$$x_1^0, x_2^0, \ldots, x_k^0$$

is called *feasible* if it can be extended to the solution of the original system (S), i.e. if there exist numbers $x_{k+1}^0, \ldots, x_n^0$ such that the set $x_1^0, \ldots, x_k^0, x_{k+1}^0, \ldots, x_n^0$ is the solution of the system (S).

Once the systems $(S'), (S'')$ and so on are constructed, we are in a position:

1) to find all feasible values of the unknown x_1 (from the system $(S^{(n-1)})$);

2) to find for any particular feasible value x_1^0 all compatible values of the unknown x_2, such that together with x_1^0 they form a feasible set (they are found by substituting x_1^0 into the system $(S^{(n-2)})$);

3) to find for any particular feasible set x_1^0, x_2^0 all compatible values of the unknown x_3 (they are found by substituting x_1^0 and x_2^0 into the system $(S^{(n-3)})$) and so on.

It is in this sense that one should understand our statement that the system (S) is solved if the systems $(S'), (S''), \ldots, (S^{(n-1)})$ are constructed.

Example. We solve in the above sense the system

$$\left.\begin{aligned} 7x + 2y - 2z - 4 &\geqslant 0 \\ -x - y - z + 4 &\geqslant 0 \\ -2x + 3y + z - 1 &\geqslant 0 \\ 5x - y + z + 2 &\geqslant 0 \end{aligned}\right\}$$

On solving each of the inequalities for z we write down the system in the form

$$\left.\begin{aligned} \tfrac{7}{2}x + y - 2 &\geqslant z \\ -x - y + 4 &\geqslant z \\ z &\geqslant 2x - 3y + 1 \\ z &\geqslant -5x + y - 2 \end{aligned}\right\} \tag{6}$$

The concomitant system is of the form

$$\left.\begin{aligned} \tfrac{7}{2}x + y - 2 &\geqslant 2x - 3y + 1 \\ \tfrac{7}{2}x + y - 2 &\geqslant -5x + y - 2 \\ -x - y + 4 &\geqslant 2x - 3y + 1 \\ -x - y + 4 &\geqslant -5x + y - 2 \end{aligned}\right\}$$

or, after the cancelling of the terms,

$$\left.\begin{aligned} \tfrac{3}{2}x + 4y - 3 &\geqslant 0 \\ \tfrac{17}{2}x &\geqslant 0 \\ -3x + 2y + 3 &\geqslant 0 \\ 4x - 2y + 6 &\geqslant 0 \end{aligned}\right\}$$

On solving each inequality for y we write down the system in the form

$$\left.\begin{aligned} y &\geqslant -\tfrac{3}{8}x + \tfrac{3}{4} \\ y &\geqslant \tfrac{3}{2}x - \tfrac{3}{2} \\ 2x + 3 &\geqslant y \\ x &\geqslant 0 \end{aligned}\right\} \tag{7}$$

The concomitant system is of the form

$$\left.\begin{aligned} 2x+3 &\geqslant -\tfrac{3}{8}x+\tfrac{3}{4} \\ 2x+3 &\geqslant \tfrac{3}{2}x-\tfrac{3}{2} \\ x &\geqslant 0 \end{aligned}\right\}$$

it is equivalent to one inequality

$$x \geqslant 0 \tag{8}$$

Thus the original system is compatible. According to the point of view we have adopted the systems (8), (7), (6) provide the solution of the set problem; namely, inequality (8) shows that there exists a solution (x, y, z) of the original system involving any nonnegative x. If a particular value of x is chosen, then from the system (7) one can find the feasible values for y. If particular values for x and y are chosen, then from the system (6) it is possible to find the feasible values of z.

Set $x = 1$, for example; then from the system (7) we have the following inequalities limiting y:

$$5 \geqslant y \geqslant \tfrac{3}{8}$$

Take, for example, $y = 4$. Setting $x = 1$, $y = 4$ in the system (6), we have the following inequalities limiting z:

$$\begin{aligned} \tfrac{11}{2} &\geqslant z \\ -1 &\geqslant z \\ z &\geqslant -9 \\ z &\geqslant -3 \end{aligned}$$

or simply

$$-1 \geqslant z \geqslant -3$$

Setting, for example, $z = -2$, we have one of the solutions of the original system: $x = 1$, $y = 4$, $z = -2$.

9. Incompatible Systems

So far we have been dealing mainly with such systems of inequalities which have at least one solution (which are compatible). As far as incompatible systems are concerned, studying them may on the face of it appear a needless pursuit; besides it seems implausible that one should be able to relate any interesting theory to such systems. In fact everything is quite different. The properties of incompatible systems are not only of interest in themselves, but also give key to the understanding of a variety of important facts. Thus the main linear programming theorem (the duality theorem, see Section 14) is in the long run derived from some properties of incompatible systems.

Consider an arbitrary system of linear inequalities. For notational convenience consider for the time being that the number of unknowns is three, although everything stated above equally applies to systems in any number of unknowns.

So we are given the system

$$\left.\begin{aligned} a_1x + b_1y + c_1z + d_1 &\geqslant 0 \\ a_2x + b_2y + c_2z + d_2 &\geqslant 0 \\ \dots\dots\dots&\dots\dots \\ a_mx + b_my + c_mz + d_m &\geqslant 0 \end{aligned}\right\} \tag{1}$$

Multiply both sides of the inequality of (1) by a nonnegative number k_1, both sides of the second inequality by a nonnegative number k_2 and so on, and then add the obtained inequalities together to come as a result to the inequality

$$\begin{aligned} &(k_1a_1 + k_2a_2 + \dots + k_ma_m)x + \\ &+ (k_1b_1 + k_2b_2 + \dots + k_mb_m)y + \\ &+ (k_1c_1 + k_2c_2 + \dots + k_mc_m)z + \\ &+ \ k_1d_1 + k_2d_2 + \dots + k_md_m \geqslant 0 \end{aligned} \tag{2}$$

which will be called a *combination of the inequalities of* (1).

It may happen that a certain combination of the inequalities of (1) is an inequality of the form

$$0 \times x + 0 \times y + 0 \times z + d \geqslant 0 \tag{3}$$

where d is a strictly negative number (division by $|d|$ then leads to the inequality $-1 \geqslant 0$). It is clear that no set of values of the unknowns satisfies such an inequality, therefore in the case under discussion the system (1) is incompatible (it has no solutions). It is quite remarkable that the converse proposition is also valid; namely, if the system (1) is incompatible, then a certain combination of its inequalities is of the form (3) where $d < 0$.

We shall now prove this proposition in the general form (i.e. for systems in any number of unknowns), but first let us introduce the following *definition*. Let us agree to say that the inequality

$$ax + by + cz + d \geqslant 0$$

is *inconsistent* if it is not satisfied by any set of values of the unknowns. Obviously, any inconsistent inequality is of the form (3) where $d < 0$ (prove this!). The proposition to be proved can now be formulated as the following theorem.

Theorem on incompatible systems of inequalities. *If a system of linear inequalities is incompatible, then a certain combination of these inequalities is an inconsistent inequality.*

The proof is made by induction on n, the number of unknowns in our system.

When $n = 1$ the system is of the form

$$\left.\begin{array}{l} a_1x + b_1 \geqslant 0 \\ a_2x + b_2 \geqslant 0 \\ \dots\dots\dots \\ a_mx + b_m \geqslant 0 \end{array}\right\} \qquad (4)$$

One can suppose that all the coefficients $a_1, a_2, \dots, a_m$ are nonzero. Indeed, if for example, $a_1 = 0$, then the first inequality is of the form $0 \times x + b_1 \geqslant 0$; if the number b_1 is nonnegative, then such an inequality may be discarded; if b_1 is negative, however, then the first inequality of our system is inconsistent and there is nothing to prove.

So let us suppose that none of the numbers $a_1, a_2, \dots, a_m$ is zero. It is easy to see that these numbers are sure to contain both positive and negative numbers. Indeed, if all of the indicated numbers had the same sign, were they positive, for example, then the system (4) would be reduced to the form

$$\left.\begin{array}{l} x \geqslant -\dfrac{b_1}{a_1} \\ x \geqslant -\dfrac{b_2}{a_2} \\ \dots\dots\dots \\ x \geqslant -\dfrac{b_m}{a_m} \end{array}\right\}$$

and would hence be compatible.

Assume for definiteness that the first k numbers of $a_1, a_2, \dots, a_k$ are positive and the remaining $m - k$ are negative. Then the system (4) is equivalent to the system

$$\left.\begin{array}{c}\boxed{\begin{array}{c} x \geqslant -\dfrac{b_1}{a_1} \\ \cdots\cdots \\ x \geqslant -\dfrac{b_k}{a_k} \end{array}} \\ \\ \boxed{\begin{array}{c} x \leqslant -\dfrac{b_{k+1}}{a_{k+1}} \\ \cdots\cdots \\ x \leqslant -\dfrac{b_m}{a_m} \end{array}}\end{array}\right\} \tag{5}$$

Choose the largest among the numbers $-(b_1/a_1), \ldots, -(b_k/a_k)$; let it be $-(b_1/a_1)$, for example. Then the first k inequalities of the system (5) can be replaced by a single (first) inequality. Similarly choose the smallest among the numbers $-(b_{k+1}/a_{k+1}), \ldots, -(b_m/a_m)$, the remaining $m-k$ inequalities of the system (5) can then be replaced by a single (last) inequality. Thus the system (4) is equivalent to the system of two inequalities:

$$\left.\begin{array}{l} x \geqslant -\dfrac{b_1}{a_1} \\ x \leqslant -\dfrac{b_m}{a_m} \end{array}\right\}$$

and its incompatibility means that

$$-\frac{b_1}{a_1} > -\frac{b_m}{a_m} \tag{6}$$

From (6) we have

$$b_m a_1 - b_1 a_m < 0 \tag{7}$$

(it should be remembered that $a_1 > 0$ and $a_m < 0$). If the first inequality of (4) is now multiplied by a positive number $-a_m$ and the last inequality by the positive number a_1, then after addition we have the inequality

$$0 \times x + (b_m a_1 - b_1 a_m) \geqslant 0 \tag{8}$$

which is by virtue of (7) inconsistent. For systems in one unknown the theorem is thus valid. We now assume that the statement of the theorem is valid for systems in $n-1$ unknowns and under this assumption establish its validity for the case of n unknowns.

So suppose an incompatible system of linear inequalities in n unknowns $x_1, x_2, \dots, x_n$ is given; keeping to the notation of the previous section we call it "system (S)". We construct the concomitant system (S') for it; the latter will be incompatible after (S). Since the number of unknowns in the system (S') is $n-1$, the hypothesis of induction applies to it. This means that a certain combination of inequalities of the system (S') is an inconsistent inequality. But it is easily seen that each inequality of the system (S') is a combination of inequalities of (S); indeed, if we simply add up the inequalities $P_\alpha \geqslant x_n$ and $x_n \geqslant Q_\beta$ of the system (S), we get $P_\alpha + x_n \geqslant x_n + Q_\beta$ or $P_\alpha \geqslant Q_\beta$, i.e. one of the inequalities of the system (S') (see Section 8, systems (T) and (S')). Hence it follows in turn that a certain combination of inequalities of the original system (S) is an inconsistent inequality. Thus the theorem is proved.

The theorem on incompatible systems of linear inequalities is just one of the manifestations of the deep analogy existing between the properties of systems of linear inequalities and those of *systems of linear equations*. Let us try to substitute the word "equation" for "inequality" in the formulation of the theorem. We shall have the following proposition:

If a system of linear equations is incompatible, then a certain combination of these equations is an inconsistent equation.

It turns out that *this proposition is also valid*. It is called the *Kronecker-Capelli theorem* in a somewhat different formulation and is proved in *linear algebra* (this is the name of the part of mathematics which studies linear operations, i.e. such operations as addition of points or multiplication of a point by a number in n-dimensional space). For the above to be better understood, however, it is necessary to introduce clarity into the concept of combination. A combination of equations is constructed in the same way as a combination of inequalities, the only difference being that one is allowed to multiply the given equations *by any*, not only nonnegative, numbers. As in the case of inequalities the word "inconsistent" refers to an equation having no solutions. It is easily shown that an inconsistent equation must certainly be of the form

$$0 \times x_1 + 0 \times x_2 + \dots + 0 \times x_n + b = 0 \tag{9}$$

where b is a nonzero number (division of both sides by b gives the "equation" $1 = 0$).

Especially important is one particular case of the theorem on incompatible systems of inequalities, namely, the case where a given system contains the inequalities

$$x_1 \geqslant 0, \quad x_2 \geqslant 0, \quad \ldots, \quad x_n \geqslant 0 \tag{10}$$

Denoting the rest of the system by (S), one can say that the problem is to find all *nonnegative* solutions of the system (S) (i.e. the solutions satisfying conditions (10)). If the problem has no solutions, then according to the theorem just proved a certain combination of inequalities of the system (S),

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + a \geqslant 0 \tag{11}$$

yields in conjunction with a certain combination of inequalities of (10),

$$k_1x_1 + k_2x_2 + \ldots + k_nx_n \geqslant 0$$
$$(k_1, k_2, \ldots, k_n \text{ are nonnegative})$$

the inconsistent inequality

$$0 \times x_1 + 0 \times x_2 + \ldots + 0 \times x_n + c \geqslant 0$$

where c is a nonnegative number. Hence

$$a_1 = -k_1 \leqslant 0,\ a_2 = -k_2 \leqslant 0,\ \ldots,\ a_n = -k_n \leqslant 0,$$
$$a < 0$$

We shall formulate the result obtained as a separate proposition.

Corollary of the theorem on incompatible systems. *If a system of inequalities has no nonnegative solutions, then a certain combination of these inequalities is an inequality of the form* (11) *where all the coefficients* $a_1, a_2, \ldots, a_n \leqslant 0$ *and the absolute term* $a < 0$.

10. A Homogeneous System of Linear Inequalities. The Fundamental Set of Solutions

In Section 8 we have discussed a method for finding solutions of systems of linear inequalities. In spite of its having many obvious merits the method fails to give answers to some questions; for example, it does not allow one to review the *set of all solutions* of a given system of inequalities. It is to this end that this and the next section of our book are devoted. The main difficulties, it will be seen, arise in considering *homogeneous* systems which are dealt with in the present section; the general case (i.e. the case of a nonhomogeneous system of inequalities) is examined in Section 11. There is no need here to confine ourselves to the case of two or three unknowns, from the outset we shall consider a system of any number of inequalities in any number of unknowns. For the reader's convenience our account is broken down into a number of points.

1°. *A linear function of n arguments.* The general form of a homogeneous inequality in n unknowns is

$$a_1x_1 + a_2x_2 + \dots + a_nx_n \geqslant 0$$

Consider separately the expression in the left-hand side of the inequality

$$a_1x_1 + a_2x_2 + \dots + a_nx_n \tag{1}$$

It is called a *linear function*. The role of the arguments is played by n variables $x_1, x_2, \dots, x_n$. One can assume, however, that function (1) depends on *one* rather than n arguments; this argument is the point

$$X = (x_1, x_2, \dots, x_n)$$

of n-dimensional space.

Let us agree from now on to designate function (1) as $L(X)$ for short:

$$L(X) = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

if we are given several functions like this instead of one, we shall designate them as $L_1(X)$, $L_2(X)$, etc.

We establish the following two properties of a linear function.

1.

$$L(kX) = kL(X)$$

where X is any point and k is any number.

2.

$$L(X + Y) = L(X) + L(Y)$$

where X and Y are any two points.

Property 1 follows in the obvious way from the equality

$$a_1(kx_1) + a_2(kx_2) + \dots + a_n(kx_n) = \\ = k(a_1x_1 + a_2x_2 + \dots + a_nx_n)$$

We now prove property 2. Let

$$X = (x_1, x_2 \dots, x_n) \quad \text{and} \quad Y = (y_1, y_2, \dots, y_n)$$

Then

$$L(X+Y) = a_1(x_1+y_1) + a_2(x_2+y_2) + \dots$$
$$\dots + a_n(x_n+y_n) = (a_1x_1 + a_2x_2 + \dots + a_nx_n) +$$
$$+ (a_1y_1 + a_2y_2 + \dots + a_ny_n) = L(X) + L(Y)$$

2°. *Some properties of the solutions of a homogeneous system of linear inequalities.* Given a homogeneous system of m linear inequalities in n unknowns

$$a_1x_1 + a_2x_2 + \dots + a_nx_n \geqslant 0 \quad \text{(the first inequality)}$$
$$b_1x_1 + b_2x_2 + \dots + b_nx_n \geqslant 0 \quad \text{(the second inequality)}$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$c_1x_1 + c_2x_2 + \dots + c_nx_n \geqslant 0 \quad \text{(the } m\text{th inequality)}$$

Designate the left-hand sides of the inequalities as $L_1(X)$, $L_2(X)$,..., $L_m(X)$ to rewrite the system as

$$\left.\begin{array}{l} L_1(X) \geqslant 0 \\ L_2(X) \geqslant 0 \\ \dots\dots\dots \\ L_m(X) \geqslant 0 \end{array}\right\} \qquad (2)$$

where

$$X = (x_1, x_2, \dots, x_n)$$

We establish the following two propositions.

1. *If the point X is a solution of the system* (2) *and k is any nonnegative number, then the point kX is again a solution of the system* (2).

2. *If X and Y are two solutions of the system* (2), *then $X + Y$ is again a solution of the system* (2).

Both propositions follow readily from the properties of a linear function proved in sect. 1°. Indeed, suppose i is any of the numbers 1, 2,..., m. We have

$$L_i(kX) = kL_i(X) \geqslant 0$$

(for $k \geqslant 0$ and $L_i(X) \geqslant 0$) as well as

$$L_i(X+Y) = L_i(X) + L_i(Y) \geqslant 0$$

(for $L_i(X) \geqslant 0$ and $L_i(Y) \geqslant 0$).

Propositions 1 and 2 immediately yield the following consequence. *If certain points X_1, X_2,..., X_p are solutions of the system* (2), *then so is any point of the form*

$$k_1X_1 + k_2X_2 + ... + k_pX_p \tag{3}$$

where k_1, k_2,..., k_p are nonnegative numbers.

Indeed, by virtue of Proposition 1 each of the points k_1X_1, k_2X_2,..., k_pX_p is a solution of the system (2) but then by virtue of Proposition 2 so is the sum of the points. Let us agree to call any point of the form (3), where k_1, k_2,..., k_p are nonnegative numbers, a *nonnegative combination* of the points X_1, X_2,..., X_p. Then the above consequence will allow the following statement.

A nonnegative combination of any number of solutions of the homogeneous system (2) *is again a solution of this system.*

3°. *The fundamental set of solutions.* We introduce the following definition.

A set of a finite number of solutions

$$X_1,\ X_2,...,\ X_p$$

of the homogeneous system (2) is said to be the *fundamental set* of solutions if any solution X of the system can be given by the formula

$$X = k_1X_1 + k_2X_2 + ... + k_pX_p \tag{4}$$

where k_1, k_2,..., k_p are nonnegative numbers. It follows that in this case formula (4) in which k_1, k_2,..., k_p are any nonnegative numbers gives a review of *all solutions* of the system (2). Hence it is clear that the problem of finding the fundamental set of solutions is one of primary importance in the investigation of the system (2). The development of an algorithm which would allow the fundamental set of solutions for any system (2) to be found using quite simple operations is what we set ourselves as the final object.

4°. *Construction of the fundamental set for a system consisting of one inequality.* Construct the fundamental set of solutions for the inequality

$$a_1x_1 + a_2x_2 + ... + a_nx_n \geqslant 0 \tag{5}$$

where the numbers a_1, a_2,..., a_n do not vanish together.

To do this, consider along with inequality (5) the equation

$$a_1x_1 + a_2x_2 + ... + a_nx_n = 0 \tag{6}$$

The properties of a linear function proved in sect. 1° readily yield the following two propositions:

1. If X is some solution of equation (6) and k is any number, then kX is again a solution of equation (6).

2. If X and Y are two solutions of equation (6), then $X+Y$ is again a solution of equation (6).

The proof of the propositions is left to the reader.

According to the premises there are some nonzero numbers among $a_1, a_2, \ldots, a_n$. Set $a_n \neq 0$, for example.

Then the equation can be written down as

$$x_n = -\frac{1}{a_n}(a_1x_1 + a_2x_2 + \ldots + a_{n-1}x_{n-1}) \tag{6'}$$

Setting $x_1 = 1$, $x_2 = 0, \ldots, x_{n-1} = 0$, we find from the last equation that $x_n = -(a_1/a_n)$. Thus the point

$$X_1 = \left(1, 0, \ldots, 0, -\frac{a_1}{a_n}\right)$$

is a solution of equation (6). Proceeding in a similar way one can obtain the solutions

$$X_2 = \left(0, 1, \ldots, 0, -\frac{a_2}{a_n}\right)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$X_{n-1} = \left(0, 0, \ldots, 1, -\frac{a_n-1}{a_n}\right)$$

Now let

$$X = (\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, \alpha_n) \tag{7}$$

be any solution of equation (6). According to (6') we have

$$\alpha_n = -\frac{1}{a_n}(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_{n-1}\alpha_{n-1}) =$$

$$= \alpha_1\left(-\frac{a_1}{a_n}\right) + \alpha_2\left(-\frac{a_1}{a_n}\right) + \ldots + \alpha_{n-1}\left(-\frac{a_{n-1}}{a_n}\right)$$

Considering the point

$$\alpha_1X_1 + \alpha_2X_2 + \ldots + \alpha_{n-1}X_{n-1} =$$

$$= \alpha_1\left(1, 0, \ldots, -\frac{a_1}{a_n}\right) + \alpha_2\left(0, 1, \ldots, -\frac{a_2}{a_n}\right) + \ldots$$

$$\ldots + \alpha_{n-1}\left(0, 0, \ldots, 1, -\frac{a_{n-1}}{a_n}\right)$$

convinces us that its coordinates coincide with those of the point X.

Therefore the following equality is valid

$$X = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_{n-1} X_{n-1} \tag{8}$$

We now add one more point

$$X_n = -(X_1 + X_2 + \dots + X_{n-1}) \tag{9}$$

to the points $X_1, X_2, \dots, X_{n-1}$ constructed earlier. It follows from the properties of the solutions of equation (6) referred to at the beginning of this subsection that the point X_n is also a solution. Now it is easy to prove the fact that *any solution X of equation* (6) *is a nonnegative combination of the solutions $X_1, X_2, \dots, X_{n-1}, X_n$.*

Indeed, let α be a positive number exceeding any of the numbers $|\alpha_1|, |\alpha_2|, \dots, |\alpha_{n-1}|$. It follows from (8) and (9) that

$$X = (\alpha_1 + \alpha) X_1 + (\alpha_2 + \alpha) X_2 + \dots$$
$$\dots + (\alpha_{n-1} + \alpha) X_{n-1} + \alpha X_n$$

which proves our statement.

For brevity of further writing denote the left-hand side of equation (6) by $L(X)$. Choose some solution of the equation $L(X) = 1$ and denote it by X_{n+1}. We state that the *set of points*

$$X_1, X_2, \dots, X_{n-1}, X_n, X_{n+1} \tag{10}$$

is the fundamental set of solutions for inequality (5).

Indeed, each of the points satisfies inequality (5). Now let X' be any solution of the inequality; hence $L(X') = a$ where $a \geqslant 0$. Then the point

$$X = X' - aX_{n+1}$$

satisfies inequality (6) for

$$L(X) = L(X') - aL(X_{n+1}) = a - a \times 1 = 0$$

If we now write down

$$X' = X + aX_{n+1}$$

and recall that the point X is a nonnegative combination of the points $X_1, \dots, X_{n-1}, X_n$ it will be clear that X' can be represented as a nonnegative combination of the points (10).

Consider a specific example. Suppose it is required to construct the fundamental set of solutions for the inequality

$$-2x_1 - 4x_2 + x_3 \geqslant 0 \tag{11}$$

in three unknowns x_1, x_2, x_3.

First of all we write down the equation

$$-2x_1 - 4x_2 + x_3 = 0$$

and solve it for one of the unknowns, x_3, for example:

$$x_3 = 2x_1 + 4x_2$$

We now successively set one of the unknowns x_1, x_2 (contained in the right-hand side of the equation) equal to 1 and the remaining unknowns to zero, and get the solutions

$$X_1 = (1,\ 0,\ 2), \quad X_2 = (0,\ 1,\ 4)$$

As X_3 we then take the point

$$X_3 = -(X_1 + X_2) = (-1,\ -1,\ -6)$$

and finally take as X_4 one of the solutions of the equation

$$-2x_1 - 4x_2 + x_3 = 1$$

$X_4 = (0,\ 0,\ 1)$, for example.

The points X_1, X_2, X_3, X_4 form the fundamental set of solutions for inequality (11). The general solution is of the form

$$\begin{aligned} X &= k_1X_1 + k_2X_2 + k_3X_3 + k_4X_4 = \\ &= k_1(1,0,2) + k_2(0,1,4) + k_3(-1,-1,-6) + k_4(0,0,1) \end{aligned}$$

or

$$\begin{aligned} x_1 &= k_1 - k_3 \\ x_2 &= k_2 - k_3 \\ x_3 &= 2k_1 + 4k_2 - 6k_3 + k_4 \end{aligned}$$

where k_1, k_2, k_3, k_4 are arbitrary nonnegative numbers.

5°. *The rearrangement of the structure of the fundamental set of solutions when another inequality is added to the system.* In order to learn how to construct fundamental sets of solutions, first consider a problem like this.

Suppose we are given the homogeneous system

$$\left.\begin{aligned} L_1(X) &\geqslant 0 \\ L_2(X) &\geqslant 0 \\ \cdots&\cdots \\ L_m(X) &\geqslant 0 \end{aligned}\right\} \qquad (12)$$

of linear inequalities. Suppose further that we know the fundamental *set of solutions*

$$X_1,\ X_2,\ \dots,\ X_p$$

for that system. It is required to construct the fundamental set of solutions for the system resulting from the addition of another inequality

$$L(X) \geqslant 0 \tag{13}$$

to (12).

The solutions of the system (12) are precisely all nonnegative combinations of the points $X_1, X_2, \dots, X_p$. We must pick among these combinations those which would satisfy inequality (13) and, what is more, constitute the fundamental set of solutions for the system (12), (13).

All points $X_1, X_2, \dots, X_p$ can be divided into three groups with respect to the function $L(X)$, the left-hand side of inequality (13): the points for which $L(X) > 0$, the points for which $L(X) < 0$, and finally the points for which $L(X) = 0$. We designate the points of the first group as $X_1^+, X_2^+, \dots, X_k^+$, the points of the second group as $X_1^-, X_2^-, \dots, X_l^-$, and those of the third group as $X_1^0, X_2^0, \dots, X_s^0$.
Thus

$$X_1^+,\ \dots,\ X_k^+,\ X_1^-,\ \dots,\ X_l^-,\ X_1^0,\ \dots,\ X_s^0$$

are the same points $X_1, X_2, \dots, X_p$ except that they are located perhaps in a different order.

All the points X_α^+ $(\alpha = 1, \dots, k)$ satisfy inequality (13) of course, and so do $X_\gamma^0 (\gamma = 1, \dots, s)$. As to the points X_β^- $(\beta = 1, \dots, l)$ none of them is a solution of inequality (13). However, of each pair

$$X_\alpha^+, \quad X_\beta^-$$

(one "plus" and one "minus" point) one can form a nonnegative combination

$$aX_\alpha^+ + bX_\beta^- \tag{14}$$

so that it should satisfy the condition $L(X) = 0$. To do this one should take

$$a = -L(X_\beta^-), \quad b = L(X_\alpha^+) \tag{15}$$

Indeed, the numbers a and b are positive and besides

$$L(aX_\alpha^+ + bX_\beta^-) = aL(X_\alpha^+) + bL(X_\beta^-) =$$
$$= -L(X_\beta^-)L(X_\alpha^+) + L(X_\alpha^+)L(X_\beta^-) = 0$$

Denote the point (14), where a and b have the values indicated above, by $X^0_{\alpha\beta}$:

$$X^0_{\alpha\beta} = -L(X^-_\beta)\,X^+_\alpha + L(X^+_\alpha)\,X^-_\beta \tag{16}$$

The answer to the set problem is given by the following Theorem. *The points*

$$X^+_1,\ \ldots,\ X^+_k,\ X^0_1,\ \ldots,\ X^0_s,\ X^0_{11},\ X^0_{12},\ \ldots,\ X^0_{kl} \tag{17}$$

(there are $k + s + kl$ points here in all) *form the fundamental set of solutions for the system* (12), (13).

To prove the theorem, we first establish the following lemma.

Lemma. *Any nonnegative combination of the points X^+_α and X^-_β can be represented as a nonnegative combination of the points X^+_α, $X^0_{\alpha\beta}$ or else as a nonnegative combination of the points X^-_β, $X^0_{\alpha\beta}$.*

Proof of the lemma. Let

$$X = cX^+_\alpha + dX^-_\beta$$

be a nonnegative combination of the points X^+_α and X^-_β. Along with X consider the point

$$X^0_{\alpha\beta} = aX^+_\alpha + bX^-_\beta$$

where the numbers a and b are given by formulas (15). Compare the two ratios, c/a and d/b. If the first is greater than the second, then, setting $d/b = \lambda$, $c/a = \lambda + \mu$ where $\mu > 0$, we have

$$X = (\lambda a + \mu a)\,X^+_\alpha + \lambda bX^-_\beta = \lambda X^0_{\alpha\beta} + \mu aX^+_\alpha$$

i.e. the point X can be represented in the form of a nonnegative combination of X^+_α and $X^0_{\alpha\beta}$. If the above ratios are equal, then $X = \lambda X^0_{\alpha\beta}$. Finally, if $c/a < d/b$, then, setting $c/a = \lambda$, $d/b = \lambda + \mu$ where $\mu > 0$, we have

$$X = \lambda X^0_{\alpha\beta} + \mu bX^-_\beta$$

Thus the lemma is proved.

Proof of the theorem. Notice first of all that each of the points (17) satisfies the system (12), (13). To prove the theorem, all that remains therefore is to check the following: if some point X is a solution of the system (12), (13), then it can be represented in the form of a nonnegative combination of the points (17).

Being a solution of the system (12), the point X can be represented as a nonnegative combination of its fundamental points $X_1, X_2, \ldots, X_p$:

$$X = a_1X^+_1 + \ldots + a_kX^+_k + b_1X^-_1 + \ldots \\ \ldots + b_lX^-_l + c_1X^0_1 + \ldots + c_sX^0_s \tag{18}$$

where all the coefficients $a_1, ..., a_k, b_1, ..., b_l, c_1, ..., c_s$ are nonnegative.

If all the numbers $b_1, ..., b_l$ are zero, then there is obviously nothing to prove. Suppose therefore that there are strictly positive numbers among those indicated above. Note that among $a_1, ..., a_k$ there are also strictly positive numbers, for otherwise we should have

$$L(X) = b_1 L(X_1^-) + ... + b_l L(X_l^-) + c_1 L(X_1^0) + ... \\ ... + c_s L(X_s^0)$$

which is impossible since X satisfies inequality (13).

Suppose for definiteness that $a_1 > 0$ and $b_1 > 0$. Making use of the lemma we can replace the sum $a_1 X_1^+ + b_1 X_1^-$ by a nonnegative combination of the points X_1^+, X_{11}^0 or of the points X_1^-, X_{11}^0. If such a replacement is made in the expression for the point X, then the total number of nonzero coefficients of $X_1^+, ..., X_k^+, X_1^-, ..., X_l^-$ will decrease at least by 1. If at the same time it is found that not all of the coefficients of $X_1^-, ..., X_l^-$ are zero in the newly obtained expression for X, then we again replace one of the sums of the form $aX_\alpha^+ + bX_\beta^-$ by a nonnegative combination of the points X_α^+, $X_{\alpha\beta}^0$ or that of the points X_β^-, $X_{\alpha\beta}^0$; as a result the number of nonzero coefficients of $X_1^+, ..., X_k^+, X_1^-, ..., X_l^-$ again decreases at least by 1. We continue in this way till an expression is obtained for the point X in which all the coefficients of $X_1^-, ..., X_l^-$ are zero. We then come to the equality of the form

$$X = a_1' X_1^+ + ... + a_k' X_k^+ + \sum_{\alpha, \beta} d_{\alpha\beta} X_{\alpha\beta}^0 + \\ + c_1 X_1^0 + ... + c_s X_s^0$$

where all the coefficients on the right are greater than or equal to zero. But this is exactly the required representation for X. Thus the theorem is proved.

6°. *The existence and the method for the construction of the fundamental set of solutions.* Let us consider an arbitrary system of homogeneous linear inequalities. For the first inequality of the system we can (using the method described in sect. 4°) construct the fundamental set of solutions. On adjoining to the first the second inequality we can, on the basis of the theorem of sect. 5°, construct the fundamental set of solutions for the system consisting of the first two inequalities. We then adjoin the third inequality and so on until we have the fundamental set of solutions for the whole of the original system of inequalities. This proves the existence, and at the same time points out the method for the construction, of the fundamental set of solutions.

Of course, if in a given system of inequalities there is a subsystem for which one can at once point out the fundamental set of solutions, then that subsystem should be taken as the starting point; successively adjoining to it the remaining inequalities one will arrive after a number of steps at the desired fundamental set.

Example. It is required to find for the system

$$\left.\begin{aligned} L_1(X) &= -3x_1 - 4x_2 + 5x_3 - 6x_4 \geqslant 0 \\ L_2(X) &= 2x_1 + 3x_2 - 3x_3 + x_4 \geqslant 0 \end{aligned}\right\} \tag{19}$$

all nonnegative solutions, i.e. all solutions which satisfy the conditions

$$\left.\begin{aligned} x_1 &\geqslant 0 \\ x_2 &\geqslant 0 \\ x_3 &\geqslant 0 \\ x_4 &\geqslant 0 \end{aligned}\right\} \tag{20}$$

To put it another way, it is required to find a general solution of the system (19), (20).

It is easily seen that for the system (20) the fundamental set of solutions will be the set of the points

$$X_1 = (1,\ 0,\ 0,\ 0), \quad X_2 = (0,\ 1,\ 0,\ 0)$$
$$X_3 = (0,\ 0,\ 1,\ 0), \quad X_4 = (0,\ 0,\ 0,\ 1)$$

(indeed, any solution $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ of the system (20) can be represented in the form $\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4$). Adjoin the first inequality of (19) to the system (20) and, using the theorem of sect. 5°, construct the fundamental set of solutions for the system obtained in this way. For computational convenience make the following table:

					$L_1(X)$
X_1	1	0	0	0	−3
X_2	0	1	0	0	−4
X_3	0	0	1	0	5
X_4	0	0	0	1	−6

Indicated in each row of the table is one of the fundamental points of the system (20) and the value of the function $L_1(X)$ for that point. It is seen from the table that the only point of X_α^+ type is X_3, that the points of X_β^- type are X_1, X_2, X_4; and that there are no points of X_γ^0 type in this case.

Find the points of $X_{\alpha\beta}^0$ type. They are

$$X_{31}^0 = 3X_3 + 5X_1 = (5,\ 0,\ 3,\ 0)$$
$$X_{32}^0 = 4X_3 + 5X_2 = (0,\ 5,\ 4,\ 0)$$
$$X_{34}^0 = 6X_3 + 5X_4 = (0,\ 0,\ 6,\ 5)$$

To avoid making further writing more complicated, designate these points as Y_1, Y_2, Y_4 (respectively) and write Y_3 instead of X_3.

The points Y_3, Y_1, Y_2, Y_4 form the fundamental set of solutions for the system consisting of (20) and the first inequality of (19).

Adjoin the second inequality of (19) to that system and make the following table:

					$L_2(Y)$
Y_3	0	0	1	0	-3
Y_1	5	0	3	0	1
Y_2	0	5	4	0	3
Y_4	0	0	6	5	-13

It is seen from the table that the role of the points Y_α^+ is now played by Y_1, Y_2 and that of the points Y_β^- by Y_3, Y_4, and that there are no points of Y_γ^0 type.

We find points of $Y_{\alpha\beta}^0$ type:

$$Y_{13}^0 = 3Y_1 + Y_3 = (15,\ 0,\ 10,\ 0) = 5(3,\ 0,\ 2,\ 0)$$
$$Y_{23}^0 = 3Y_2 + 3Y_3 = (0,\ 15,\ 15,\ 0) = 5(0,\ 3,\ 3,\ 0)$$
$$Y_{14}^0 = 13Y_1 + Y_4 = (65,\ 0,\ 45,\ 5) = 5(13,\ 0,\ 9,\ 1)$$
$$Y_{24}^0 = 13Y_2 + 3Y_4 = (0,\ 65,\ 70,\ 15) = 5(0,\ 13,\ 14,\ 3)$$

The points Y_1, Y_2, Y_{13}^0; Y_{23}^0, Y_{14}^0, Y_{24}^0 form the fundamental set of solutions for the system (19), (20). The general solution is of the form

$$X = aY_1 + bY_2 + cY_{13}^0 + dY_{23}^0 + eY_{14}^0 + fY_{24}^0 =$$
$$= a(5,\ 0,\ 3,\ 0) + b(0,\ 5,\ 4,\ 0) + 5c(3,\ 0,\ 2,\ 0) +$$
$$+ 5d(0,\ 3,\ 3,\ 0) + 5e(13,\ 0,\ 9,\ 1) + 5f(0,\ 13,\ 14,\ 3)$$

where a, b, c, d, e, f are any nonnegative numbers.

Setting $a = k_1$, $b = k_2$, $5c = k_3$, $5d = k_4$, $5e = k_5$, $5f = k_6$, we have for X the representation

$$X = k_1(5,\ 0,\ 3,\ 0) + k_2(0,\ 5,\ 4,\ 0) + k_3(3,\ 0,\ 2,\ 0) +$$
$$+ k_4(0,\ 3,\ 3,\ 0) + k_5(13,\ 0,\ 9,\ 1) + k_6(0,\ 13,\ 14,\ 3) \quad (21)$$

where k_1, $k_2, \ldots, k_6$ are any nonnegative numbers.

To conclude this point, we note that having proved the existence of the fundamental set of solutions for any homogeneous system of linear inequalities we have thereby proved Theorem 2′ of Section 7 on the structure of any convex polyhedral cone.

11. The Solution of a Nonhomogeneous System of Inequalities

Now that we have learnt how to construct the general solution of a homogeneous system of inequalities it won't be difficult to solve a similar problem for an arbitrary, i.e. nonhomogeneous, system of inequalities.

Let

$$\left.\begin{array}{l} a_1x_1 + a_2x_2 + \ldots + a_nx_n \geqslant a \\ b_1x_1 + b_2x_2 + \ldots + b_nx_n \geqslant b \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ c_1x_1 + c_2x_2 + \ldots + c_nx_n \geqslant c \end{array}\right\} \quad (1)$$

be an arbitrary system of linear inequalities in n unknowns x_1, $x_2, \ldots, x_n$. Along with it consider the homogeneous system

$$\left.\begin{array}{l} a_1x_1 + a_2x_2 + \ldots + a_nx_n - at \geqslant 0 \\ b_1x_1 + b_2x_2 + \ldots + b_nx_n - bt \geqslant 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_1x_1 + c_2x_2 + \ldots + c_nx_n - ct \geqslant 0 \end{array}\right\} \quad (2)$$

in $n + 1$ unknowns x_1, $x_2, \ldots, x_n$, t.

There is a relation of certain kind between solutions of the systems

(1) and (2) in that if $x_1, x_2, \ldots, x_n, t$ is a solution of the system (2), with $t > 0$, then the numbers

$$\tilde{x}_1 = \frac{x_1}{t}, \quad \tilde{x}_2 = \frac{x_2}{t}, \quad \ldots, \quad \tilde{x}_n = \frac{x_n}{t} \tag{3}$$

will constitute a solution of the system (1).

Indeed, consider, for example, the first inequality of (2). It holds for the numbers $x_1, x_2, \ldots, x_n, t$, i.e.

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n \geqslant at$$

On dividing both sides of the inequality by a positive t we have

$$a_1\tilde{x}_1 + a_2\tilde{x}_2 + \ldots + a_n\tilde{x}_n \geqslant a$$

but this means that the set of the numbers $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ satisfies the first inequality of (1). A similar argument goes through for any inequality of (2), whence our statement.

It is easily seen that any solution of the system (1) can be obtained by the above method. Indeed, if $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ is some solution of the system (1), then the numbers $x_1 = \tilde{x}_1$, $x_2 = \tilde{x}_2, \ldots, x_n = \tilde{x}_n$, $t = 1$ satisfy the system (2) and at the same time equalities (3) are valid.

Thus, in order to find all the solutions $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ of the system (1) one should find all the solutions of the system (2) for which $t > 0$, and transform each of them according to formulas (3).

Example. Suppose it is required to find all the solutions of the following system

$$\left.\begin{aligned} -3x_1 - 4x_2 + 5x_3 &\geqslant 6 \\ 2x_1 + 3x_2 - 3x_3 &\geqslant -1 \\ x_1 &\geqslant 0 \\ x_2 &\geqslant 0 \\ x_3 &\geqslant 0 \end{aligned}\right\} \tag{4}$$

Proceeding as pointed out above, write down the following auxiliary homogeneous system

$$\left.\begin{aligned} -3x_1 - 4x_2 + 5x_3 - 6t &\geqslant 0 \\ 2x_1 + 3x_2 - 3x_3 + t &\geqslant 0 \\ x_1 &\geqslant 0 \\ x_2 &\geqslant 0 \\ x_3 &\geqslant 0 \end{aligned}\right\}$$

Adjoin to this system the inequality $t \geqslant 0$ (for we are interested only in such solutions for which $t > 0$) to get the system (19), (20) of Section 10, the only difference being that we write t instead of x_4. The solution set of this system is, as shown in Section 10 (see (21) of Section 10), given by the formulas

$$\begin{aligned} x_1 &= 5k_1 + 3k_3 + 13k_5 \\ x_2 &= 5k_2 + 3k_4 + 13k_6 \\ x_3 &= 3k_1 + 4k_2 + 2k_3 + 3k_4 + 9k_5 + 14k_6 \\ t &= k_5 + 3k_6 \end{aligned}$$

where $k_1, k_2, \ldots, k_6$ are any nonnegative numbers. Since we are interested only in solutions for which $t > 0$, it should be assumed that at least one of the numbers k_5, k_6 is nonzero (strictly positive). We now find the general solution of the system (4) from the formulas

$$\left.\begin{aligned} x_1 &= \frac{5k_1 + 3k_3 + 13k_5}{k_5 + 3k_6} \\ x_2 &= \frac{5k_2 + 3k_4 + 13k_6}{k_5 + 3k_6} \\ x_3 &= \frac{3k_1 + 4k_2 + 2k_3 + 3k_4 + 9k_5 + 14k_6}{k_5 + 3k_6} \end{aligned}\right\} \qquad (5)$$

It should be emphasized once again that in these formulas $k_1, k_2, \ldots, k_6$ are any nonnegative numbers, at least one of the numbers k_5, k_6 being other than zero.

Having established that solving the system (1) by the above method reduces to solving the homogeneous system (2) we have proved Theorem 3 of Section 7 on the structure of any convex polyhedral region. This may be illustrated by the example of the system (4). Set

$$k_i' = \frac{k_i}{k_5 + 3k_6} \quad (i = 1, 2, 3, 4)$$

since the numbers k_1, k_2, k_3, k_4 are arbitrary nonnegative, so are the numbers k_1', k_2', k_3', k_4'. Further set

$$k_5' = \frac{k_5}{k_5 + 3k_6}, \quad k_6' = \frac{3k_6}{k_5 + 3k_6}$$

the numbers k_5' and k_6' are nonnegative and constrained by the condition $k_5' + k_6' = 1$. Formulas (5) can now be written down in the form of a single equality

$$\begin{aligned} (x_1, x_2, x_3) = k_1'(5, 0, 3) + k_2'(0, 5, 4) + k_3'(3, 0, 2) + \\ + k_4'(0, 3, 3) + k_5'(13, 0, 9) + k_6'\left(0, \frac{13}{3}, \frac{14}{3}\right) \end{aligned} \qquad (6)$$

We introduce the notation:

$$X_1 = (5,\ 0,\ 3), \quad X_2 = (0,\ 5,\ 4), \quad X_3 = (3,\ 0,\ 2)$$

$$X_4 = (0,\ 3,\ 3), \quad X_5 = (13,\ 0,\ 9), \quad X_6 = \left(0,\ \frac{13}{3},\ \frac{14}{3}\right)$$

Bearing in mind the above restrictions on k_1', k_2', k_3', k_4' as well as on k_5', k_6', equality (6) can now be interpreted as follows: the solution set of the system (4) is $(X_1, X_2, X_3, X_4) + \langle X_5, X_6 \rangle$. Thereby the statement of Theorem 3 of Section 7 is proved for the system (4).

12. A Linear Programming Problem

Linear programming is a relatively new field of applied mathematics which developed in the 1940s or 1950s in connection with the tackling of varied economic problems.

As a rule problems occurring in economics, and in production planning in particular, are problems of finding the *most profitable variant*. Each of them calls for an answer to the question as to how to use the limited available resources to obtain the utmost effect. Until recently the only method of solving such problems has been by ordinary rough calculation, estimating "by sight", or else by looking all the possible variants over to find the best. The situation is changing now. Over the past few decades the complexity of production has increased to such an extent that it has become impossible just to look the variants over. The factors influencing decisions have turned out to be so numerous that in some cases the number of variants is running into milliards. This has drastically increased the interest in mathematical methods in economics. Besides, the process of "mathematization of economics" has been promoted by the development of computing technique, in particular by the advent of electronic computers.

Let us consider some examples of linear programming problems.

Example 1. A factory turning out items of two types has an assembly shop with a production capacity of 100 items of the first type or 300 items of the second per diem; at the same time the quality control department is capable of checking at most 150 items (of either type) per diem. It is further known that items of the first type cost twice as much as items of the second type. Under these conditions it is required to find such a production plan (so many items of the first type and so many items of the second per diem) which would ensure the largest profit for the factory.

The desired production plan is specified by means of two nonnegative integers x, y (x being the number of items of the first type and y that of the second) which must satisfy the following conditions*:

$$3x + y \leqslant 300; \quad x + y \leqslant 150$$

$$2x + y \quad \text{is maximal}$$

In other words, from among the nonnegative integer solutions of the system

$$\left.\begin{aligned} 3x + y \leqslant 300 \\ x + y \leqslant 150 \end{aligned}\right\} \qquad (1)$$

one should pick the one imparting the largest value to the linear function

$$f = 2x + y$$

Given an xOy rectangular coordinate system, the solution set of the system (1) will be represented by the shaded polygon of Fig. 47.

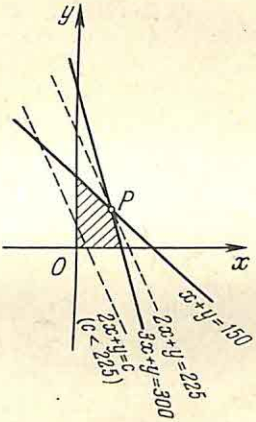


Fig. 47

---

* The first condition comes from the assembly shop. Indeed, it can turn out three items of the second type instead of one of the first. Hence, in terms of second-type items the shop's entire production is $3x + y$ articles; this number must not exceed 300.

From the same figure one can see that the point P (75, 75), one of the vertices of the polygon, is the solution of the problem.

Indeed, consider a straight line $2x + y = c$ where c is a certain number. Denote the line by l_c. As the number c increases, the line l_c shifts "upwards" (at the same time remaining parallel to its initial position). The largest value of c at which the line l_c still has points in common with the shaded polygon is that value of c at which the line passes through the point P. Hence at this point the function $2x + y$ attains its largest value (in comparison with its values at the rest of the points of the polygon).

The example discussed is certainly very primitive, still it gives an idea of the nature of linear programming problems. In all of them it is required to find the maximal (or minimal) value of some linear function of n variables

$$f = c_1x_1 + c_2x_2 + ... + c_nx_n$$

on condition that all these variables obey a given system of linear inequalities (which must without fail contain the nonnegativity constraints on the variables: $x_1 \geqslant 0, ..., x_n \geqslant 0$). In some problems (such as in the example above) an additional requirement is imposed on the variables, that they should be integers.

Example 2 (*diet problem*). There are several kinds of food at one's disposal. It is necessary to compose a diet which would, on the one hand, satisfy the minimal human body requirements for nutrients (proteins, fats, carbohydrates, vitamins, etc.) and involve the smallest expenditures, on the other.

Consider a simple mathematical model of this problem.

Suppose there are two kinds of food, F_1 and F_2, which contain the nutrients A, B, C. It is known how much nutrient of one or another kind is contained in 1 lb of F_1 or F_2; this information is presented in the following table.

	A	B	C
F_1, 1 lb	a_1	b_1	c_1
F_2, 1 lb	a_2	b_2	c_2

In addition to these data we know that a, b, c are daily requirements of human body for A, B, C (respectively) and that s_1, s_2 are costs of 1 lb of F_1 and F_2 (respectively). It is necessary to compute the amount x_1 of the food F_1 and the amount x_2 of the food F_2 so that the required amounts of the nutrients should be ensured with minimal expenditures on the food.

Obviously the total cost of the food will be

$$S = s_1 x_1 + s_2 x_2$$

The total amount of the nutrient A in both kinds of food is $a_1 x_1 + a_2 x_2$. It must not be less than a:

$$a_1 x_1 + a_2 x_2 \geqslant a$$

Similar inequalities must hold for B and C: $b_1 x_1 + b_2 x_2 \geqslant b$, $c_1 x_1 + c_2 x_2 \geqslant c$. Thus we arrive at the following problem.

We are given the system

$$\left.\begin{aligned} a_1 x_1 + a_2 x_2 \geqslant a \\ b_1 x_1 + b_2 x_2 \geqslant b \\ c_1 x_1 + c_2 x_2 \geqslant c \end{aligned}\right\} \qquad (2)$$

of three linear inequalities in two unknowns x_1, x_2 and the linear function

$$S = s_1 x_1 + s_2 x_2$$

It is required to pick from among the nonnegative solutions (x_1, x_2) of system (2) such a solution that the function S attains the smallest value (is minimized). Various problems can be reduced to such schemes, viz. alloy problems and problems in fuel-oil blending, feed mix problems and problems in fertilizer mixing, etc.

Example 3 (*transportation problem*). Coal mined in several deposits is shipped to a number of consumers, to factories, power stations, etc. It is known how much coal is mined at each of the deposits, say, per month and how much is required by any of the consumers for the same term. One knows the distances between the deposits and the consumers as well as the traffic conditions between them; considering these data one can calculate the cost of transporting each ton of coal from any deposit to any consumer. It is required that the transportation of coal should be planned in such a way under these conditions that the expenditures on it shall be minimal.

Assume for simplicity that there are only two deposits D_1, D_2 and three consumers C_1, C_2, C_3. The amounts of coal at D_1 and D_2 are a_1 and a_2 respectively; let the requirements of C_1, C_2, C_3 be b_1, b_2, b_3 respectively. Let us suppose that the total stocks of coal are equal to the total requirements:

$$a_1 + a_2 = b_1 + b_2 + b_3$$

such an assumption is quite natural. Finally, the numbers c_{ij} $(i = 1, 2; j = 1, 2, 3)$ are given, denoting the costs of transporting a ton of coal from D_i to C_j. The problem is to find the six numbers

$$x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}$$

where x_{ij} is the amount of coal due to be shipped from D_i to C_j.

For the convenience of reviewing we make the following table:

	To C_1	To C_2	To C_3	Amount shipped
From D_1	x_{11}	x_{12}	x_{13}	a_1
From D_2	x_{21}	x_{22}	x_{23}	a_2
Amount delivered	b_1	b_2	b_3	

The total amount of coal shipped from D_1 must be a_1; hence we have the condition

$$x_{11} + x_{12} + x_{13} = a_1$$

A similar condition must hold for D_2:

$$x_{21} + x_{22} + x_{23} = a_2$$

The total amount of coal delivered to C_1 must be b_1; hence

$$x_{11} + x_{21} = b_1$$

Similarly we get the conditions

$$x_{12} + x_{22} = b_2, \quad x_{13} + x_{23} = b_3$$

We assume that the cost of transportation is directly proportional to the amount of coal conveyed, i.e. transportation of coal from D_i to C_j costs $c_{ij}x_{ij}$. The total cost of transportation will then be

$$S = c_{11}x_{11} + c_{12}x_{12} + c_{13}x_{13} + c_{21}x_{21} + c_{22}x_{22} + c_{23}x_{23}$$

Thus we arrive at the following problem.

Given the system

$$\left.\begin{aligned} x_{11}+x_{12}+x_{13}&=a_1\\ x_{21}+x_{22}+x_{23}&=a_2\\ x_{11}+x_{21}&=b_1\\ x_{12}+x_{22}&=b_2\\ x_{13}+x_{23}&=b_3 \end{aligned}\right\} \qquad (3)$$

of five linear equations in six unknowns and the linear function S, it is required to pick from among the nonnegative solutions

$$x_{11},\ x_{12},\ x_{13},\ x_{21},\ x_{22},\ x_{23}$$

of the system (3) such a solution that the function S attains its smallest value (is minimized).

The problem can certainly be formulated in a more general form, i.e. with any number of deposits and consumers. It was given the name of "transportation problem" and was one of the first to be successfully solved using linear programming methods.

We have considered in a simplified form three linear programming problems. Despite the diversity of their subjects the problems have much in common in their formulations. In each problem one seeks the values of several unknowns, it being required:

(i) that these values should be nonnegative;

(ii) that these values should satisfy a certain system of linear equations or linear inequalities;

(iii) that at these values a certain linear function should attain a minimum (or maximum).

We remark that condition (ii) seems to be a disjoining rather than a uniting condition, for in one case the unknowns must satisfy equations and in the other they must satisfy inequalities. But we shall see later that one case is easily reduced to the other.

It is problems of this kind that linear programming is concerned with. More strictly, *linear programming is a branch of mathematics which studies methods of finding the minimal (or maximal) value of a linear function of several variables provided these satisfy a finite number of linear equations and inequalities*. The number of variables and that of conditions (equations or inequalities) may of course be arbitrary. In actual problems these numbers may be very large (some ten or several tens or even more).

Let us put the above verbal statement down in strict terms of formulas. The general mathematical statement of a linear programming problem looks as follows.

Given a system of linear equations

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array}\right\} \tag{I}$$

and a linear function

$$f = c_1x_1 + c_2x_2 + \ldots + c_nx_n \tag{II}$$

Find a nonnegative solution

$$x_1 \geqslant 0,\ x_2 \geqslant 0,\ \ldots,\ x_n \geqslant 0 \tag{III}$$

of the system (I) *such that the function* f *assumes the minimal value* (*is minimized*).

The equations of (I) are called the *constraints* of a given problem. Strictly speaking, conditions (III) are also constraints; it is not customary, however, to call them so since they are not characteristic of a *given* problem alone, but are common to all linear programming problems.

Of the three examples considered above only ~~the~~ last (the transportation problem) corresponds, one would think, to the statement just given. The remaining two look slightly different since all the constraints they contain have the form of inequalities and not of equations. It is possible, however, using a simple method, to make inequality constraints go over into the equivalent constraints given in the form of equations.

Assume, indeed, that the constraints of a given problem contain the inequality

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b \geqslant 0 \tag{4}$$

We introduce a new, so-called *additional*, unknown x_{n+1} connected with the unknowns $x_1, x_2, \ldots, x_n$ by the equation

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b = x_{n+1}$$

Inequality (4) is obviously equivalent to the nonnegativity condition on x_{n+1}. If an additional unknown is introduced for each of the inequalities contained in the system of constraints for a given problem, requiring in addition that all additional unknowns should be nonnegative, the problem will assume standard form (I), (II), (III) although involving a larger number of unknowns. We demonstrate this method using the diet problem as an example. The

system of constraints consists of three inequalities in this problem, so we introduce three additional unknowns x_3, x_4, x_5.
As a result the constraints assume the form of equations:

$$\left.\begin{aligned} a_1x_1 + a_2x_2 - x_3 = a \\ b_1x_1 + b_2x_2 - x_4 = b \\ c_1x_1 + c_2x_2 - x_5 = c \end{aligned}\right\}$$

It is required to find among all the nonnegative solutions of this system such a solution which minimizes the function

$$S = s_1x_1 + s_2x_2$$

Certainly, what we are essentially concerned with in the required solution are the values of x_1 and x_2.

It is frequent for a linear programming problem to require that the maximum, and not the minimum, of a linear function f should be found out. Since $\max f = -\min(-f)$ (establish this for yourself), one problem is reduced to the other if f is replaced by $-f$.

We have shown that a linear programming problem having inequality constraints can be reduced to a problem with equation constraints. It is highly interesting that the converse should also be possible, i.e. any linear programming problem can be stated in such a way that all constraints will assume the form of inequalities. We shall not dwell on this, however.

In conclusion we shall make a few remarks concerning the adopted terminology. Any nonnegative solution of a system of constraints is called *feasible*. A feasible solution yielding the minimum of a function f is called *optimal*. It is the finding of an optimal solution that is our goal. The optimal solution, if any, is not necessarily unique; cases are possible where there may be an infinite number of optimal solutions.

13. The Simplex Method

Actual linear programming problems tend to contain a large number of constraints and unknowns. It is natural that solving such problems involves a large number of calculations. This difficulty is overcome with the help of high-speed computers. The algorithm a computer programme is based on may be connected with a specific class of problems. Thus, for example, there are very simple algorithms for solving the transportation problem which are conditioned by the peculiarities of its system of constraints. There exist, however, general methods as well which allow a solution of a linear programming problem to be found in a visible number of steps. The first and foremost of them is the so-called *simplex method* with some of its modifications*.

1°. *A description of the simplex method.* So let us consider a linear programming problem. Given a certain system of linear equations in n unknowns $x_1, x_2, \ldots, x_n$ and a certain linear function f, it is required to find among the nonnegative solutions of the given system such a solution which minimizes the function f.

To begin working by the simplex method it is necessary that the given system of equations should be reduced to a form such that some r unknowns are expressed in terms of the rest, the *absolute terms in the expressions being nonnegative.* Assume, for example, that $n=5$ and that the unknowns expressed in terms of the rest are x_1, x_2, x_3. Hence the system of constraints is reduced to the form

$$\left.\begin{aligned} x_1 &= \alpha + \alpha_4 x_4 + \alpha_5 x_5 \\ x_2 &= \beta + \beta_4 x_4 + \beta_5 x_5 \\ x_3 &= \gamma + \gamma_4 x_4 + \gamma_5 x_5 \end{aligned}\right\} \tag{1}$$

where

$$\alpha \geqslant 0, \quad \beta \geqslant 0, \quad \gamma \geqslant 0 \tag{2}$$

Is it really possible to reduce the system to such a form and how is this to be done? This question will be considered later on (see sect. 2°). The unknowns x_1, x_2, x_3 in the left members of the system (1) are called *basic* and the entire set $\{x_1, x_2, x_3\}$ which for brevity we shall denote by a single letter B is called a *basis*; the remaining unknowns are called *nonbasic* or *free***. Substituting into the original expression for the function f the expressions in terms of the nonbasic unknowns of (1) for the basic unknowns, we can write down the function f itself in the nonbasic unknowns x_4, x_5:

$$f = c + c_4 x_4 + c_5 x_5$$

We set the nonbasic unknowns equal to zero

$$x_4 = 0, \quad x_5 = 0$$

---

* The name of the simplex method has nothing to do with the essence of the matter and is due to a casual circumstance.

** The latter name is due to the fact that when finding the solutions of the system (1) (irrespective of a linear programming problem) these unknowns may be assigned any values.

and find from the system (1) the values of the basic unknowns:

$$x_1 = \alpha, \quad x_2 = \beta, \quad x_3 = \gamma$$

The resulting solution of the system

$$(\alpha, \beta, \gamma, 0, 0)$$

will, due to (2), be feasible. It is called a *basic* solution corresponding to the basis $B = \{x_1, x_2, x_3\}$. For the basic solution the value of the function f is

$$f_B = c$$

The solution of a linear programming problem by the simplex method falls into a number of steps. Each step consists in going over from a given basis B to another basis B' so that the value of f should decrease or at least not increase: $f_{B'} \leqslant f_B$. The new basis B' is derived from the old basis B in a very simple way: one of the unknowns is eliminated from B and another (a former nonbasic) unknown introduced instead. A change in the basis involves a corresponding change in the structure of the system (1), of course. After a certain number k of such steps we either come to a basis $B^{(k)}$ for which $f_{B(k)}$ is the desired minimum of the function f and the corresponding basic solution is optimal, or else find out that the problem has no solution.

We illustrate the simplex method by some examples.

Example 1. Let a given system of constraints and a function f be reduced to the following form:

$$\left.\begin{aligned} x_1 &= 1 - x_4 + 2x_5 \\ x_2 &= 2 + 2x_4 - x_5 \\ x_3 &= 3 - 3x_4 - x_5 \end{aligned}\right\}$$

$$f = x_4 - x_5$$

Here the unknowns x_1, x_2, x_3 form the basis. The corresponding basic solution is

$$(1, 2, 3, 0, 0)$$

the value of f is 0 for this solution.

We find out whether the solution is optimal. Since x_5 has in f a negative coefficient, we may try to decrease the value of f by increasing x_5 (while retaining a zero value for x_4). Care should be taken in doing so, however, since changing x_5 will lead to changes in the values of x_1, x_2, x_3 and it is necessary to see to it that none of them become negative.

There is no such danger for x_1, since increasing x_5 results in the increase of x_1. Examining x_2 and x_3 we find that x_5 can be increased to 2 (at most, otherwise x_2 will become negative) which yields $x_1 = 5$, $x_2 = 0$, $x_3 = 1$. As a result we have a new feasible solution (5, 0, 1, 0, 2) in which the number of positive unknowns is equal to three as before. The value of f for this solution is equal to -2.

The new basis now consists of x_1, x_5, x_3. To make appropriate modifications in the system of constraints, it is necessary to express these unknowns in terms of x_2, x_4. We begin with the equation for x_2 (the new nonbasic unknown) solving it for x_5 (the new basic unknown):

$$x_5 = 2 + 2x_4 - x_2$$

we then express x_1, x_3 and f:

$$\begin{aligned} x_1 &= 1 - x_4 + 2(2 + 2x_4 - x_2) \\ x_3 &= 3 - 3x_4 - (2 + 2x_4 - x_2) \\ f &= x_4 - (2 + 2x_4 - x_2) \end{aligned}$$

Thus the problem is reduced to the form

$$\left.\begin{aligned} x_1 &= 5 + 3x_4 - 2x_2 \\ x_5 &= 2 + 2x_4 - x_2 \\ x_3 &= 1 - 5x_4 + x_2 \end{aligned}\right\}$$
$$f = -2 - x_4 + x_2$$

The new basic solution is

$$(5,\ 0,\ 1,\ 0,\ 2)$$

the value of f for this solution is equal to -2. This completes the first step of the process.

Let us see if it is possible to decrease the value of f still further. The coefficient of x_4 in the expression for f is negative, therefore it is possible to try to decrease f by increasing x_4 (without changing $x_2 = 0$). There is nothing in the first and the second equation to prevent it, and it is seen from the third that x_4 can be increased to 1/5 (at most, for otherwise x_3 will become negative). Setting $x_4 = 1/5$, $x_2 = 0$, we have $x_1 = 28/5$, $x_5 = 12/5$, $x_3 = 0$. As a result we arrive at a new feasible solution (28/5, 0, 0, 1/5, 12/5).

The new basis now consists of x_1, x_5, x_4. The equation for x_3 (the new nonbasic unknown) is solved for x_4 (the new basic unknown):

$$x_4 = \frac{1}{5} - \frac{1}{5}x_3 + \frac{1}{5}x_2$$

and the resulting expression is substituted into the other equations. As a result the system assumes the form

$$\left.\begin{aligned} x_1 &= \frac{28}{5} - \frac{3}{5}x_3 - \frac{7}{5}x_2 \\ x_5 &= \frac{12}{5} - \frac{2}{5}x_3 - \frac{3}{5}x_2 \\ x_4 &= \frac{1}{5} - \frac{1}{5}x_3 + \frac{1}{5}x_2 \end{aligned}\right\}$$

and the following expression is obtained for the function f:

$$f = -\frac{11}{5} + \frac{1}{5}x_3 + \frac{4}{5}x_2$$

The new basic solution is

$$(28/5,\ 0,\ 0,\ 1/5,\ 12/5)$$

the corresponding value of f being equal to $-11/5$. This completes the second step of the process.

Since in the last expression for the function f both unknowns x_3 and x_2 have positive coefficients, the minimum of f is attained when $x_3 = x_2 = 0$. This means that the last basic solution (28/5, 0, 0, 1/5, 12/5) is optimal and the required minimum f is equal to $-11/5$; thus the problem is solved.

In the example analysed the process has terminated in finding an optimal solution. Yet another termination of the process is possible. To illustrate it, we solve the following example.

Example 2.

$$\left.\begin{aligned} x_3 &= 1 + x_1 - x_2 \\ x_4 &= 2 - x_1 + 2x_2 \\ f &= -x_1 - x_2 \end{aligned}\right\}$$

Here the basis is formed by the unknowns x_3, x_4. The basic solution is of the form

$$(0,\ 0,\ 1,\ 2)$$

the corresponding value of the function f being equal to 0.

In the expression for f the coefficient of x_1 is negative, therefore we try to increase x_1 (without changing $x_2 = 0$). The first equation does not prevent us from doing so, and it is seen from the second equation that x_1 can be increased only to 2, which yields $x_3 = 3$, $x_4 = 0$. The new feasible solution is (2, 0, 3, 0); for this solution the value of f is equal to -2.

The new basis now consists of x_3, x_1. We solve the second equation for x_1 and substitute the resulting expression into the first equation and f. The problem assumes the following form

$$\left.\begin{aligned} x_3 &= 3 - x_4 + x_2 \\ x_1 &= 2 - x_4 + 2x_2 \end{aligned}\right\}$$
$$f = -2 + x_4 - 3x_2$$

The new basic solution is

$$(2,\ 0,\ 3,\ 0)$$

the corresponding value of f being equal to -2.

In the last expression for f the coefficient of x_2 is negative. We want to see how much we can decrease the value of f by increasing x_2 (while keeping $x_4 = 0$). For this purpose we look over the x_2 terms in both equations of the system and notice that both coefficients of x_2 are positive. Thus x_2 can be increased without limit while keeping x_3 and x_1 positive, f taking arbitrarily large absolute negative values. So min $f = -\infty$ (it is said that the function f is not bounded from below) and there exists no optimal solution.

As is seen from the above examples, each successive step in the simplex method is based on the choice of a negative coefficient for some nonbasic unknown x_j in the expression for f. If it turns out that there are several negative coefficients, one may choose any of them. This introduces a certain amount of arbitrariness into the computations. The arbitrariness does not necessarily end here, however; it may happen that increasing x_j to the extreme makes several basic unknowns vanish together. Then it is possible to make any of them x_i exchange roles with x_j, i.e. to make x_i nonbasic and, on the contrary, introduce x_j into the basis.

It would of course be possible to eliminate the arbitrariness by introducing some additional agreement. There is no particular need to do it, however. In fact some arbitrariness is even helpful, for it varies the computation process and hence allows one to seek such a sequence of steps which would lead as quickly as possible to the solution of the problem.

2°. *Finding the first basis.* In the previous subsection we described the procedure for solving a linear programming problem by the simplex method. We required as a preliminary condition that the system of constraints should be reduced to the form *

$$\left.\begin{aligned} x_1 &= \alpha + \alpha_{r+1}x_{r+1} + \dots + \alpha_n x_n \\ x_2 &= \beta + \beta_{r+1}x_{r+1} + \dots + \beta_n x_n \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ x_r &= \gamma + \gamma_{r+1}x_{r+1} + \dots + \gamma_n x_n \end{aligned}\right\} \qquad (3)$$

where $\alpha \geqslant 0$, $\beta \geqslant 0$, ..., $\gamma \geqslant 0$; then we say that the unknowns x_1, x_2, ..., x_r form the basis.

In many linear programming problems the basis can be directly perceived, and in others it has to be found. We shall consider one of the methods of finding the basis generally known as "the method of an artificial basis".

We shall analyse an example to illustrate the method.

Example. Given the system of constraints

$$\left.\begin{aligned} 2x_1 - 2x_2 + 2x_3 - x_4 - 7x_5 &= -4 \\ -x_1 + 3x_2 - 2x_3 + x_4 + 3x_5 &= 5 \end{aligned}\right\}$$

it is required to solve it for a certain original basis.

Let us first of all transform the system so that the absolute terms of the equations should be nonnegative. To do this we should multiply both members of the first equation by -1. We get the following system

$$\left.\begin{aligned} -2x_1 + 2x_2 - 2x_3 + x_4 + 7x_5 &= 4 \\ -x_1 + 3x_2 - 2x_3 + x_4 + 3x_5 &= 5 \end{aligned}\right\} \qquad (4)$$

We now introduce auxiliary, or *artificial*, unknowns y_1, y_2 (one for each equality) in the following way:

$$\left.\begin{aligned} y_1 &= 4 - (-2x_1 + 2x_2 - 2x_3 + x_4 + 7x_5) \\ y_2 &= 5 - (-x_1 + 3x_2 - 2x_3 + x_4 + 3x_5) \end{aligned}\right\} \qquad (5)$$

Obviously all solutions of the original system (4) can be obtained by taking all the solutions of the system (5) which satisfy the conditions $y_1 = 0$, $y_2 = 0$, and keeping only the values of $x_1, \dots, x_5$ in each of those solutions.

The unknowns y_1, y_2 form the basis in the system (5). Suppose that starting from this basis we succeed in going over to another basis which contains no artificial unknowns. Then omitting the terms containing y_1 and y_2 from the resulting equations (or, equivalently, setting $y_1 = 0$, $y_2 = 0$) we get a system equivalent to the original system (4), this system being solved for a certain basis.

---

* Or to a similar form where the left members contain some other r unknowns rather than x_1, $x_2, \dots, x_r$; the remaining $n - r$ unknowns must enter into the right members.

It remains to decide how to transform the unknowns y_1, y_2 of the system (5) into nonbasic unknowns. It is remarkable that we should be able to use the simplex method for this purpose too; namely, using this method we shall work out the problem of minimizing the function

$$F = y_1 + y_2$$

under constraints (5) and the conditions $x_1 \geqslant 0, \ldots, x_5 \geqslant 0, y_1 \geqslant 0, y_2 \geqslant 0$. Here the premises under which the simplex method begins to work are satisfied since the system (5) is of the required form (it has the original basis y_1, y_2). In some steps we arrive at the desired minimum. Since $F \geqslant 0$ and hence $\min F \geqslant 0$, two cases are possible:

1. $\min F > 0$. This means that the system (5) has no nonnegative solutions for which $y_1 = 0$, $y_2 = 0$ (otherwise $\min F = 0$). Therefore the *original system* (4) *will have no nonnegative solutions in this case*. Hence it will follow that any linear programming problem with such a system of constraints is insoluble.

2. $\min F = 0$. Suppose that in this case $(x_1^0, \ldots, x_5^0, y_1^0, y_2^0)$ is an optimal solution. Since $y_1^0 + y_2^0 = \min F = 0$, we have $y_1^0 = 0$, $y_2^0 = 0$. It follows that $(x_1^0, \ldots, x_5^0)$ will be a nonnegative solution of the original system (4).

Thus *in the case of* $\min F = 0$ *the system of constraints* (4) *will have at least one nonnegative solution.*

Further, if on accomplishing the process (i.e. when $\min F = 0$ has been attained) it is found that the artificial unknowns y_1, y_2 are among the nonbasic ones, we have achieved our object — we have separated the basis from the unknowns $x_1, \ldots, x_5$. If, however, it becomes clear that some of the unknowns y_1, y_2 still belong to the basis, then some further transformations are required.

So we solve the problem of minimizing the function

$$\begin{aligned} F = y_1 + y_2 = [4 - (-2x_1 + 2x_2 - 2x_3 + x_4 + 7x_5)] + \\ + [5 - (-x_1 + 3x_2 - 2x_3 + x_4 + 3x_5)] = \\ = 9 + 3x_1 - 5x_2 + 4x_3 - 2x_4 - 10x_5 \end{aligned}$$

In the expression for F the coefficient of x_4 is negative, we therefore try to decrease the value of F by increasing x_4. Comparing the coefficients of x_4 and the absolute terms in the equations of (5) we find that x_4 can be increased only to 4, y_1 then vanishing. The new basis consists of x_4, y_2. We solve the first equation for x_4 and substitute the resulting expression into the second equation and into F. The problem takes the form

$$\left.\begin{aligned} x_4 &= 4 + 2x_1 - 2x_2 + 2x_3 - 7x_5 - y_1 \\ y_2 &= 1 - x_1 - x_2 \qquad\quad + 4x_5 + y_1 \end{aligned}\right\} \qquad (6)$$
$$F = 1 - x_1 - x_2 \qquad\quad + 4x_5 + 2y_1$$

This completes the first step of the process. As a result we have succeeded in making the unknown y_1 nonbasic.

In the new expression for F the coefficient of x_1 is negative. Looking at equations (6) we see that x_1 can be increased only to 1, y_2 then vanishing. The new basis consists of x_4, x_1. On transforming equations (6) and the expression for F (transforming the latter is now optimal, however) we have

$$\left.\begin{aligned} x_4 &= 6 - 4x_2 + 2x_3 + x_5 + y_1 - 2y_2 \\ x_1 &= 1 - x_2 \qquad\quad + 4x_5 + y_1 - y_2 \\ F &= y_1 + y_2 \end{aligned}\right\} \qquad (7)$$

Thus as a result of the second step the artificial unknowns y_1, y_2 have become nonbasic. Eliminating from equations (7) the terms having y_1, y_2, we get

$$\left.\begin{aligned} x_4 &= 6 - 4x_2 + 2x_3 + x_5 \\ x_1 &= 1 - x_2 \qquad\quad + 4x_5 \end{aligned}\right\}$$

This means that we have separated a basis in the original system (4). Thus the problem is solved.

As already noted, when applying the simplex method to finding the minimum of the function F equal to the sum of the artificial unknowns, it may happen that by the time the process is over (i.e. when $\min F = 0$ has already been attained) some of the artificial unknowns still make part of the basis. These unknowns can also be transformed into nonbasic ones by simple methods which we shall demonstrate using an example.

Suppose, for example, that after a certain number of steps the problem has assumed the following form

$$\left.\begin{aligned} x_2 &= \frac{1}{2} - \frac{1}{2}x_1 + \frac{3}{2}x_3 - \frac{3}{2}x_4 - \frac{1}{2}x_5 - \frac{1}{2}y_1 \\ y_2 &= x_1 - x_3 + 11x_4 + 5x_5 + 2y_1 \\ y_3 &= 6x_1 + 6x_3 + 6x_4 + 4x_5 + 2y_1 \\ F &= 7x_1 + 5x_3 + 17x_4 + 9x_5 + 5y_1 \end{aligned}\right\} \qquad (8)$$

We see that the function F has already attained the minimum equal to zero, it may therefore be omitted from further considerations. The unknowns y_2 and y_3 are still part of the basis. But the absolute terms in the equations for y_2 and y_3 are zero and not by chance, for min $F = 0$ implies that in the resulting basic solution $y_2 = 0$ and $y_3 = 0$. We now make use of this circumstance to make y_2 and y_3 leave the basis. To do this we observe that there is a negative number among the coefficients of the unknowns $x_1, \ldots, x_5$ in the equation for y_2: it is the coefficient -1 of x_3. We make the replacement $x_3 \leftrightarrow y_2$ (we make the unknown x_3 basic and make y_2 become nonbasic). To do this we solve the second of the equations for x_3 and substitute the resulting expression for x_3 into the other equations. Since the absolute term in the second equation is zero, as a result of such an operation the absolute terms in the equations will remain unchanged (and hence nonnegative).

We get:

$$\left.\begin{aligned} x_2 &= \frac{1}{2} + x_1 + 15x_4 + 7x_5 + \frac{5}{2}y_1 - \frac{3}{2}y_2 \\ x_3 &= x_1 + 11x_4 + 5x_5 + 2y_1 - y_2 \\ y_3 &= 12x_1 + 72x_4 + 34x_5 + 14y_1 - 6y_2 \end{aligned}\right\} \quad (9)$$

Now we have no negative coefficients (of $x_1, \ldots, x_5$) in the equation for y_3, therefore we shall not be able to make y_3 nonbasic. But we should not be much annoyed at this. For if some numbers $x_1, \ldots, x_5$ satisfy the original system of constraints, then y_1, y_2 and y_3 must be zero. Thus

$$12x_1 + 72x_4 + 34x_5 = 0$$

Since here all the coefficients of the unknowns have the same sign, this must yield

$$x_1 = x_4 = x_5 = 0$$

these equalities are thus consequences of the original system of constraints and the conditions $x_i \geqslant 0$ $(i = 1, \ldots, 5)$. If we regard them as satisfied the system (9) reduces to

$$\left.\begin{aligned} x_2 &= \frac{1}{2} \\ x_3 &= 0 \end{aligned}\right\}$$

So subject to the nonnegativity of the unknowns the given system of constraints allows only one solution:

$$x_2 = \frac{1}{2}, \quad x_3 = 0, \quad x_1 = 0, \quad x_4 = 0, \quad x_5 = 0$$

14. The Duality Theorem in Linear Programming

There are so-called *duality theorems* in various branches of mathematics. Each of them allows the constructing for any statement in a given theory, according to a certain standard rule, of another statement in such a way that the validity of the first statement automatically leads to the validity of the second. A remarkable example of a duality theorem is met in linear programming too.

Suppose a certain linear programming problem is given; we shall call it the *original problem*, or *problem* A, below. The constraints are given in the form of *inequalities*. So we have the system

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_1 \geqslant 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_2 \geqslant 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n + b_m \geqslant 0 \end{aligned}\right\} \qquad (1)$$

of m linear inequalities in n unknowns and a linear function

$$f = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

We must pick from all nonnegative solutions ($x_1 \geqslant 0$, $x_2 \geqslant 0, \dots, x_n \geqslant 0$) such that the function f attains the *largest* value (f is maximized).

We connect with problem A another problem which we call a *dual problem* relative to problem A or *problem* A′. It is stated as follows: given the system

$$\left.\begin{aligned} a_{11}y_1 + a_{21}y_2 + \dots + a_{m1}y_m + c_1 \leqslant 0 \\ a_{12}y_1 + a_{22}y_2 + \dots + a_{m2}y_m + c_2 \leqslant 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{1n}y_1 + a_{2n}y_2 + \dots + a_{mn}y_m + c_n \leqslant 0 \end{aligned}\right\} \qquad (1')$$

of n linear inequalities in m unknowns and the linear function

$$\varphi = b_1y_1 + b_2y_2 + \dots + b_my_m$$

It is necessary to pick all nonnegative solutions of the system (1′) such that the function φ attains the *smallest* value possible (φ

is minimized). Comparing problems A and A′ we notice:

1. That the coefficient of the jth unknown in the ith equation of the system (1′) is the same as that of the ith unknowns in the jth equation of the system (1);

2. That the absolute terms of the inequalities in each of the problems coincide with the coefficients of the unknowns in the linear function of the other problem;

3. That in the system of inequalities of problem A the inequalities are all of the $\geqslant 0$ type, it being required in the problem that the maximum of f should be attained. On the contrary, in the system of inequalities of problem A′ the inequalities are all of the $\leqslant 0$ type, but in return it is required in it that the minimum of φ should be attained.

One of the main theorems in linear programming, the so-called *duality theorem*, states the following.

Duality theorem. *If an original problem is solvable, so is its dual, the maximum of the function f being equal to the minimum of the function φ:*

$$max\ f = min\,\varphi$$

We shall prove this theorem by reducing it to the question of compatibility of a certain system of inequalities.

To make the proof more convenient to follow, we break it down into several stages.

Stage 1. Lemma. *If $x_1^0, \ldots, x_n^0$ is some nonnegative solution of the system (1) and $y_1^0, \ldots, y_m^0$ is some nonnegative solution of the system (1′), then for these solutions the values of the functions f and φ are connected by the inequality*

$$f_0 \leqslant \varphi_0$$

Proof. We consider the inequalities of the system (1) where the values $x_1^0, \ldots, x_n^0$ are substituted for $x_1, \ldots, x_n$. We multiply the first of the inequalities by y_1^0, the second by y_2^0, etc., and then add all the inequalities obtained:

$$(a_{11}y_1^0x_1^0 + \ldots + a_{mn}y_m^0x_n^0) + b_1y_1^0 + \ldots + b_my_m^0 \geqslant 0$$

(one should bear in mind that we are multiplying the inequalities by nonnegative numbers, therefore the signs of the inequalities remain unchanged). In the same way, we multiply the first inequality of the system (1′) by x_1^0, the second by x_2^0, etc., and then add the resulting inequalities together:

$$(a_{11}y_1^0x_1^0 + \ldots + a_{mn}y_m^0x_n^0) + c_1x_1^0 + \ldots + c_nx_n^0 \leqslant 0$$

In both cases the brackets contain an expression equal to a sum of terms of the form $a_{ij}y_i^0x_j^0$ over all $i=1,\dots, m$, $j=1,\dots, n$. Hence the two expressions in brackets coincide. But then

$$c_1x_1^0 + \dots + c_nx_n^0 \leqslant b_1y_1^0 + \dots + b_my_m^0$$

or $f_0 \leqslant \varphi_0$. Thus the lemma is proved.

Stage 2. *Reducing problems* A *and* A′ *to the solution of a certain system of inequalities.*

Consider the following "combined" system of inequalities:

$$\left.\begin{array}{l|r}
a_{11}x_1 + \dots + a_{1n}x_n & +b_1 \geqslant 0 \\
\dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\
a_{m1}x_1 + \dots + a_{mn}x_n & +b_m \geqslant 0 \\
\hline
 & a_{11}y_1 + \dots + a_{m1}y_m + c_1 \leqslant 0 \\
\dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\
 & a_{1n}y_1 + \dots + a_{mn}y_m + c_n \leqslant 0 \\
\hline
c_1x_1 + \dots + c_nx_n & -b_1y_1 - \dots - b_my_m \quad \geqslant 0
\end{array}\right\} \qquad (S)$$

It is seen that it is made up of the system (1), the system (1′) and the inequality $f-\varphi \geqslant 0$. The unknowns of the system (S) are $x_1,\dots, x_n, y_1,\dots, y_m$ (there are $n+m$ unknowns in all). We first of all establish the following fact.

If the system (S) *has nonnegative solution* $x_1^0,\dots, x_n^0, y_1^0,\dots, y_m^0$, *then the numbers* $x_1^0,\dots, x_n^0$ *give a solution to problem* A *and the numbers* y_1^0, y_m^0 *a solution to problem* A′, *with* $f_0 = \varphi_0$.

We shall linger on this proposition a little longer to emphasize the principal role it plays. What is remarkable about it is that a *linear programming problem, i.e. a maximization problem, reduces to solving a certain system of linear inequalities without any maximization requirements.* In fact, the solution of the system (S) (in the range of nonnegative values of unknowns) is of course not a bit easier than the solution of the original linear programming problem (problem A); however, the very possibility of such a reduction is very curious.

Now we prove the above statement. First of all, it is clear that the numbers $x_1^0,\dots, x_n^0$ are nonnegative and satisfy the system (1); similarly the numbers $y_1^0,\dots, y_m^0$ are nonnegative and satisfy (1′). Moreover, these numbers satisfy the inequality

$$f_0 \geqslant \varphi_0$$

(following from the last inequality of the system (S)). On the other hand, by lemma we have

$$f_0 \leqslant \varphi_0$$

Hence, $f_0 = \varphi_0$.

Further, if $x_1, \dots, x_n$ is *some* nonnegative solution of the system (1), then by the lemma we again have

$$f \leqslant \varphi_0$$

Comparing this and $f_0 = \varphi_0$ we get $f \leqslant f_0$, whence it follows that f_0 is the maximal value of f.

Similarly, if $y_1, \dots, y_m$ is some nonnegative solution of the system (1′), then by the lemma we have

$$f_0 \leqslant \varphi$$

comparing this and $f_0 = \varphi_0$ we get $\varphi_0 \leqslant \varphi$, i.e. φ_0 is the minimal value of φ. This proves the above proposition.

Stage 3. *Completing the proof.* Now it remains for us to show the following: if problem A has a solution, then the system (S) has a nonnegative solution. For then, as shown above, $f_0 = \varphi_0$, i.e. $\max f = \min \varphi$.

We shall argue following the rule of contraries, i.e. we shall suppose that the system (S) has no nonnegative solutions. For this case we have the corollary of the theorem on incompatible systems (Section 9). True, this corollary refers to a system consisting of inequalities of the $\geqslant 0$ type and our system (S) has inequalities of the $\leqslant 0$ type. But this is easy to mend by writing the system (S) in the form

$$\left.\begin{array}{l|r} a_{11}x_1 + \dots + a_{1n}x_n & + b_1 \geqslant 0 \\ \dots\dots\dots\dots & \dots\dots\dots\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n & + b_m \geqslant 0 \\ \hline & -a_{11}y_1 - \dots - a_{m1}y_m - c_1 \geqslant 0 \\ \dots\dots\dots\dots & \dots\dots\dots\dots \\ & -a_{1n}y_1 - \dots - a_{mn}y_m - c_n \geqslant 0 \\ \hline c_1x_1 + \dots + c_nx_n & -b_1y_1 - \dots - b_my_m \geqslant 0 \end{array}\right\} \qquad (S')$$

So suppose that the system (S') has no nonnegative solutions. According to the corollary of the theorem on incompatible systems, there will occur nonnegative numbers $k_1, \dots, k_m$, $l_1, \dots, l_n$, s ($m + n + 1$

numbers all in all) such that*

$$\left.\begin{array}{l} a_{11}k_1 + \ldots + a_{m1}k_m + c_1 s \leqslant 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}k_1 + \ldots + a_{mn}k_m + c_n s \leqslant 0 \end{array}\right\} \qquad (2)$$

$$\left.\begin{array}{l} -a_{11}l_1 - \ldots - a_{1n}l_n - b_1 s \leqslant 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ -a_{m1}l_1 - \ldots - a_{mn}l_n - b_m s \leqslant 0 \end{array}\right\} \qquad (2')$$

$$b_1k_1 + \ldots + b_mk_m - c_1l_1 - \ldots - c_nl_n < 0 \qquad (3)$$

We first of all show that the number s is nonzero. Indeed, suppose this is not the case, i.e. that $s=0$. Consider some nonnegative solution $x_1^0, \ldots, x_n^0$ of the system (1). Arguing in the same way as we did in proving the lemma, we find that

$$(a_{11}k_1x_1^0 + \ldots + a_{mn}k_mx_n^0) + b_1k_1 + \ldots + b_mk_m \geqslant 0$$

But the parenthesis $\leqslant 0$ (this results if we multiply the first inequality of (2) by x_1^0, the second by x_2^0 and so on, bearing in mind $s=0$, and then carry out addition). Hence

$$b_1k_1 + \ldots + b_mk_m \geqslant 0 \qquad (4)$$

after which it follows from (3) that

$$c_1l_1 + \ldots + c_nl_n > 0$$

Further, by multiplying inequalities (2′) (where $s = 0$) by an arbitrary positive number λ we get

$$\begin{array}{l} a_{11}\lambda l_1 + \ldots + a_{1n}\lambda l_n \geqslant 0 \\ \cdots\cdots\cdots\cdots\cdots \\ a_{m1}\lambda l_1 + \ldots + a_{mn}\lambda l_n \geqslant 0 \end{array} \qquad (5)$$

If we now add to each inequality of (1), where $x_1, \ldots, x_n$ are substituted for by $x_1^0, \ldots, x_n^0$, the corresponding inequality of (5), we find that the numbers

$$x_1^0 + \lambda l_1, \ldots, x_n^0 + \lambda l_n$$

---

* The numbers $k_1, \ldots, k_m, l_1, \ldots, l_n, s$ are precisely those numbers by which we multiply the first, second, ..., $(m+n+1)$th inequalities of the system (S') to get (after addition) the inconsistent inequality $a_1x_1 + \ldots + a_nx_n + a'_1y_1 + \ldots + a'_my_m + d \geqslant 0$, where $a_1, \ldots, a_n, a'_1, \ldots, a'_m \leqslant 0$ and d is a negative number.

also constitute a nonnegative solution of the system (1). For this solution the value of the function f is

$$c_1x_1^0 + \dots + c_nx_n^0 + \lambda(c_1l_1 + \dots + c_nl_n)$$

and since the parenthesis is strictly positive, the value of f increases indefinitely with λ. This means that $\max f = \infty$, i.e. that problem A, contrary to the premises, has no solution.

Thus s is nonzero. It then follows from (2) that the numbers $k_1/s, \dots, k_m/s$ are a nonnegative solution of the system (1), from (2′) that the numbers $l_1/s, \dots, l_n/s$ are a nonnegative solution of the system (1′), and from (3) that for these solutions $\varphi - f < 0$. But this is contrary to the lemma. So on supposing that the system (S) has no nonnegative solutions, we come to a contradiction. Such solutions are therefore certain to exist, which proves the duality theorem.

Example. Find the maximal value of the function

$$f = 2x_2 + 12x_3$$

provided the variables x_1, x_2, x_3 are nonnegative and satisfy the inequalities

$$\left.\begin{aligned} x_1 - x_2 - x_3 + 2 &\geqslant 0 \\ -x_1 - x_2 - 4x_3 + 1 &\geqslant 0 \end{aligned}\right\}$$

Solution. Let us call the set problem problem A. The dual problem (problem A′) must be stated as follows: find the minimal value of the function

$$\varphi = 2y_1 + y_2$$

provided the variables y_1, y_2 are nonnegative and satisfy the inequalities

$$\left.\begin{aligned} y_1 - y_2 &\leqslant 0 \\ -y_1 - y_2 + 2 &\leqslant 0 \\ -y_1 - 4y_2 + 12 &\leqslant 0 \end{aligned}\right\} \qquad (6)$$

Problem A′ can be solved graphically by representing in the y_1Oy_2 coordinate plane the feasible region of the system (6). This is done in Fig. 48. It is seen from the same figure that the

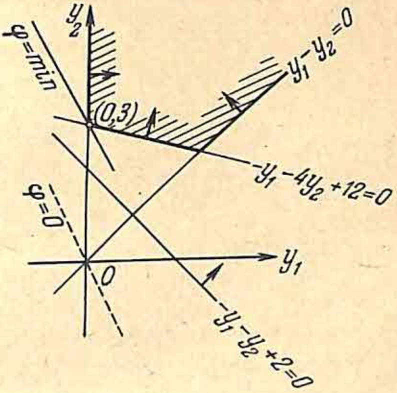


Fig. 48

function φ attains its smallest value at the point (0,3), one of the vertices of the region. This value is equal to 3. By the duality theorem the maximum of the function f must also be equal to 3.

15. Transportation Problem

In Section 12 we have considered a number of specific linear programming problems. Of special interest among them is the transportation problem, above all due to its practical significance. There is voluminous literature devoted to this and similar problems. One must say that the methods of solving the transportation problem are rather instructive. They demonstrate the indubitable fact that when solving any special class of problems general methods (say the simplex method) should be employed with caution, i.e. one should consider in full measure the peculiarities of the given class of problems.

1°. *Stating the problem.* We remind the reader how the transportation problem is stated in the general form. There are some, say m, source points (suppliers)

$$A_1, A_2, \ldots, A_m$$

and so many, say n, destination points (consumers)

$$B_1, B_2, \ldots, B_n$$

We use a_i to denote the quantity of goods (say in tons) concentrated at point $A_i (i = 1, 2, \ldots, m)$ and b_j to denote the quantity of goods expected at point B_j ($j = 1, 2, \ldots, n$). We assume the condition

$$a_1 + a_2 + \ldots + a_m = b_1 + b_2 + \ldots + b_n$$

implying that the total stock of goods is equal to the summed demand for it. In addition to the numbers a_i, b_j, we are also given quantities c_{ij} denoting the cost of transporting a ton of goods from point A_i to point B_j. It is required to develop an optimal shipping schedule, i.e. to compute what quantity of goods must be shipped from each source point to each destination point for the total cost of shipments to be minimal.

So our problem has mn nonnegative unknown numbers x_{ij} ($i = 1, \ldots, m$; $j = 1, \ldots, n$) where x_{ij} is the quantity of goods to be shipped from A_i to B_j. Lest we should complicate the things with cumbersome expressions, we take specific values for m and n, namely, $m = 3$ and $n = 4$. So we shall assume that there are three source points, A_1, A_2, A_3, and four destination points, B_1, B_2, B_3, B_4. The unknown quantities x_{ij} can be summarized in a table called the *transportation table*:

Table 1

Destination point / *Source point*	*To* B_1	*To* B_2.	*To* B_3	*To* B_4	*Stocks*
From A_1	x_{11}	x_{12}	x_{13}	x_{14}	a_1
From A_2	x_{21}	x_{22}	x_{23}	x_{24}	a_2
From A_3	x_{31}	x_{32}	x_{33}	x_{34}	a_3
Requirements	b_1	b_2	b_3	b_4	

Since the total weight of the goods shipped from A_1 to all consumption points must be equal to a_1, we have the relation

$$x_{11} + x_{12} + x_{13} + x_{14} = a_1$$

Similar relations must hold for the points A_2 and A_3. The unknowns of our problem must therefore satisfy the equations

$$\left.\begin{aligned} x_{11} + x_{12} + x_{13} + x_{14} &= a_1 \\ x_{21} + x_{22} + x_{23} + x_{24} &= a_2 \\ x_{31} + x_{32} + x_{33} + x_{34} &= a_3 \end{aligned}\right\} \qquad (1)$$

But we should also bear in mind that the total quantity of goods delivered to B_1 from all source points must be equal to b_1, i.e.

$$x_{11}+x_{21}+x_{31}=b_1$$

Similar relations must hold for the points B_2, B_3, B_4. This leads to the equations

$$\left.\begin{aligned} x_{11}+x_{21}+x_{31}&=b_1\\ x_{12}+x_{22}+x_{32}&=b_2\\ x_{13}+x_{23}+x_{33}&=b_3\\ x_{14}+x_{24}+x_{34}&=b_4 \end{aligned}\right\} \tag{2}$$

Notice that equations (1) and (2) are very easy to remember. Indeed, the ith equation of the system (1) implies that the sum of the unknowns in the ith row of the transportation table is equal to a_i; we shall call equations (1) *horizontal* for this reason. Similarly, the jth equation of the system (2) records the fact that the sum of the unknowns in the jth column of the transportation table is equal to b_j; because of this we shall call the equations of the system (2) *vertical.*

We ship x_{ij} tons of goods from point A_i to point B_j, the cost of transporting one ton being c_{ij}. Hence transportation from A_i to B_j costs $c_{ij}x_{ij}$ and the total cost of transportation will be

$$S=\sum_{i,j} c_{ij}x_{ij} \tag{3}$$

where the symbol $\sum\limits_{i,j}$ put before $c_{ij}x_{ij}$ signifies that the quantities $c_{ij}x_{ij}$ must be summed over all $i = 1, 2, 3$, and all $j = 1, 2, 3, 4$ (there will be 12 summands in all).

We thus come to the following linear programming problem.

Given the system of equations (1), (2) *and the linear function* (3). *Find among the nonnegative solutions of the system such that minimizes the function* (3).

The transportation problem can be solved by the simplex method like any other linear programming problem. Because of the special structure of the system of constraints (1), (2) the general procedure of the simplex method is, of course, greatly simplified when applied to the transportation problem. Here we present a method for solving the transportation problem called the *method of potentials*. It is a variant of the simplex method specially adapted for solving the transportation problem.

2°. *Finding the first basis*. As we know, the work by the simplex method is preceded by a preparatory stage, that of finding the first basis. In the case of the transportation problem there is a very simple and convenient method of finding the first basis called the *northwest corner rule*. The essence of the method is best explained by considering a specific example. Suppose we are given three source points, A_1, A_2, A_3, and four destination points, B_1, B_2, B_3, B_4, the stocks and requirements being equal to the following quantities:

$$a_1 = 60, \quad a_2 = 80, \quad a_3 = 100$$
$$b_1 = 40, \quad b_2 = 60, \quad b_3 = 80, \quad b_4 = 60$$

These data are tabulated in Table 2 below.

Table 2

	B_1	B_2	B_3	B_4	*Stocks*
A_1	40				60
A_2					80
A_3					100
Requirements	40	60	80	60	

Let us try to satisfy the requirements of the first destination point B_1 using the stocks of the first source point A_1. This can be done in this case since the stocks $a_1 = 60$ are greater than the requirements $b_1 = 40$.

We therefore enter the number 40 in the cell x_{11}. The requirements of the point b_1 will be found fully satisfied, and therefore the B_1 column may be temporarily excluded from consideration. It is now possible to consider that we have come to new Table 3 which has three destination points, B_2, B_3, B_4, and three source points, A_1, A_2, A_3, the stocks at A_1 being $a'_1 = 60 - 40 = 20$. Note that in Table 3 the sum of all requirements is as before equal to the sum of all stocks.

We apply the same method to Table 3 and try to satisfy the requirements $b_2 = 60$ of the point B_2 (it assumes the role of the first point in Table 3) using the stocks $a'_1 = 20$ of the point A_1. It is obvious that we can satisfy these requirements only partially since $b_2 > a_1'$.

Table 3

	B_2	B_3	B_4	
A_1	20			20
A_2				80
A_3				100
	60	80	60	

We enter the number 20, this being the maximum of what can be transported from A_1 to B_2, in the cell x_{12}. The requirements of B_2 will be reduced to $b_2' = 40$ and the stocks of A_1 will be found fully exhausted. In virtue of this the A_1 column of Table 3 may be temporarily removed. Thus we come to Table 4 which has only two source points, A_2 and A_3, and three destination points, B_2, B_3, B_4, now.

Table 4

	B_2	B_3	B_4	
A_2	40			80
A_3				100
	40	80	60	

In a similar way we continue to reduce successively the resulting tables until we satisfy the requirements of all the destination points. In virtue of the condition of the problem all the stocks of the source points will be found exhausted*.

We get the following values for some of the unknowns in the course of reducing the tables:

---

* When "new" stocks are equal to "new" requirements, either the row or the column may be excluded at will from the table.

$$x_{11}=40, \quad x_{12}=20, \quad x_{22}=40, \quad x_{23}=40, \quad x_{33}=40, \quad x_{34}=60 \tag{4}$$

On entering them in Table 2 we get Table 5.

Table 5

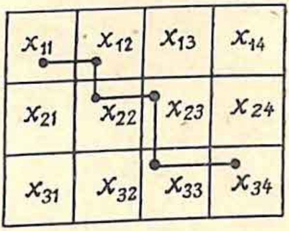
	B_1	B_2	B_3	B_4
A_1	40	20		
A_2		40	40	
A_3			40	60

Let us agree to refer to those cells in Table 5 which have the values of unknowns entered in them as *basic*, and to the remaining cells as *vacant*. If we consider that the values of the unknowns x_{ij} which correspond to the vacant cells are zero, then the resulting set of all the unknowns yields the feasible solution of our problem. Indeed, it is easy to verify that the sum of the values of the unknowns in each row of the table is equal to the stocks of the corresponding source point and that in each column, to the requirements of the corresponding destination point. Therefore equations (1), (2) are satisfied. It now remains to remark that the values of all unknowns are nonnegative.

In the general case, i.e. when there are any number m of source points and any number n of destination points, the method we have described allows us to fill $m+n-1$ cell in the table. Indeed, in every step we fill exactly one cell, after which one row or one column is deleted from the table; there is one exception, where the table consists of a single cell on filling which we exclude at once both the row and the column. Since the number of all rows is m and the number of all columns is n, it is clear that the number of all steps and hence the number of filled (basic) cells is $m+n-1$.

We now show that the system of constraints (1), (2) can be solved for the unknowns corresponding to the basic cells. Hence it will follow that these unknowns may be assumed to be basic and the remaining unknowns corresponding to the vacant cells in the table to be nonbasic.

To prove this, we join the basic cells with a broken line in the order they arose in the procedure described above. There will result a broken line of the following form for the example considered:



It is not difficult to see now that the unknowns occupying the basic cells can be expressed in terms of the unknowns occupying the vacant cells. The expressions are found in consecutive order:

first for x_{11}, taking the first vertical equation;
then for x_{12}, taking the first horizontal equation;
then for x_{22}, taking the second vertical equation;
then for x_{23}, taking the second horizontal equation;
then for x_{33}, taking the third vertical equation;
finally for x_{34}, taking the third horizontal (or the fourth vertical) equation.

So we have found the basic set for our problem:

$$x_{11},\ x_{12},\ x_{22},\ x_{23},\ x_{33},\ x_{34}$$

This does not complete the solution of the set problem since the expressions for the basic unknowns in terms of the nonbasic ones still remain to be found. But these expressions will not be required in explicit form at the next stages of solving the transportation problem.

3°. *Solving the problem by the method of potentials.* The finding of the first basis is but a preparatory stage of solving the problem. When this stage is over, all the unknowns are found broken down into two groups:

x_{kl}, basic unknowns, and x_{pq}, nonbasic unknowns

No actual expressions for the basic unknowns in terms of nonbasic ones will be needed. As to the function S (the total cost of transportation), it is obligatory to know its expression in terms of nonbasic unknowns. In other words, it is necessary to express the function S as

$$S = \sum_{p,q} s_{pq} x_{pq} + s \tag{5}$$

We shall show how to find coefficients s_{pq} of nonbasic unknowns.

We assign to each source point A_i a certain value α_i ($i = 1, 2, \dots, m$), the "potential" of point A_i. Similarly, to each of the destination points B_j we assign a value β_j ($j = 1, 2, \dots, n$), the "potential" of point B_j. We relate these values as follows: for every basic unknown x_{kl} we work out an equation

$$\alpha_k + \beta_l = c_{kl} \tag{6}$$

where c_{kl} is used as before to denote the cost of transporting a ton of goods from a point A_k to a point B_l. The set of the equations of the form (6) worked out for all basic unknowns x_{kl} forms a system of linear equations. This system has $m + n - 1$ equations (as many as there are basic unknowns) and $m + n$ unknowns α_k, β_l (as many as there are source points and destination points taken together). It can be shown that this system is always compatible, it being possible to arbitrarily specify the value of one of the unknowns and then unambiguously find the values of the remaining unknowns from the system.

Let us fix some one solution ($\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n$) of the system of equations (6) and then calculate the sum $\alpha_p + \beta_q$ for each nonbasic unknown x_{pq}. We denote this sum by c'_{pq}:

$$\alpha_p + \beta_q = c'_{pq}$$

and call it *indirect cost* (in contrast to the real cost c_{pq}). It then turns out that in the expression (5) that interests us the coefficients of nonbasic unknowns are equal to

$$s_{pq} = c_{pq} - c'_{pq} \tag{7}$$

Formula (7) will be proved in the next section.

If all the variables s_{pq} are nonnegative, then the initial basic solution will be optimal. If, however, there are negative variables, say $s_{p_0q_0}$, among them, then we go over to the next basis. It will be recalled that this step begins with reasoning as follows. We shall increase $x_{p_0q_0}$ (keeping the other nonbasic unknowns zero). If in doing so we reach a moment when one of the basic unknowns, say $x_{k_0l_0}$, vanishes, we go over to a new basis by eliminating the unknown $x_{k_0l_0}$ from the old basis and introducing instead $x_{p_0q_0}$. Going over to the new basis completes one step of the simplex method.

So far we have presented only some of the considerations necessary for the understanding of the method of potentials. As to the actual solution of the problem, it can be best explained by considering a specific example.

Example. We are given three source points, A_1, A_2, A_3, with the stocks

$$a_1 = 60, \quad a_2 = 80, \quad a_3 = 100$$

and four destination points, B_1, B_2, B_3, B_4, having the requirements

$$b_1 = 40, \quad b_2 = 60, \quad b_3 = 80, \quad b_4 = 60$$

The values c_{ij} (costs of transportation from A_i to B_j per ton of goods) are given by Table 6.

Table 6

1	2	3	4
4	3	2	0
0	2	2	1

Solution. We begin with the finding of the first basic solution. It this case the northwest corner rule leads to the result shown in Table 7 (cf. the example of the previous subsection). We denote this solution by $\underset{1}{X}$ for short. The value of the function S is as follows for it:

Table 7

40	20		
	40	40	
		40	60

$$\Sigma c_{ij}x_{ij} = 1 \times 40 + 2 \times 20 + 3 \times 0 + 4 \times 0 + 4 \times 0 + 3 \times 40 +$$

$$+ 2 \times 40 + 0 \times 0 + 0 \times 0 + 2 \times 0 + 2 \times 40 + 1 \times 60 = 420$$

To find the potentials it is necessary to solve the system of equations

$$\left.\begin{aligned}\alpha_1 + \beta_1 &= c_{11} = 1\\ \alpha_1 + \beta_2 &= c_{12} = 2\\ \alpha_2 + \beta_2 &= c_{22} = 3\\ \alpha_2 + \beta_3 &= c_{23} = 2\\ \alpha_3 + \beta_3 &= c_{33} = 2\\ \alpha_3 + \beta_4 &= c_{34} = 1\end{aligned}\right\}$$

any solution of this system suiting us. As already noted, the value of one of the unknowns can be specified arbitrarily. Setting $\alpha_1 = 1$, we find successively:

$$\beta_1 = 0, \quad \beta_2 = 1, \quad \alpha_2 = 2, \quad \beta_3 = 0, \quad \alpha_3 = 2, \quad \beta_4 = -1$$

We then compute the indirect costs c'_{pq}:

$$\begin{aligned}c'_{13} &= \alpha_1 + \beta_3 = 1'\\ c'_{14} &= \alpha_1 + \beta_4 = 0'\\ c'_{21} &= \alpha_2 + \beta_1 = 2'\\ c'_{24} &= \alpha_2 + \beta_4 = 1'\\ c'_{31} &= \alpha_3 + \beta_1 = 2'\\ c'_{32} &= \alpha_3 + \beta_2 = 3'\end{aligned}$$

It is more convenient to carry out the above computation using Table 8 in which we first enter only the values of c_{kl} for all the basic cells. Assuming $\alpha_1 = 1$ and proceeding according to the above rule we compute the potentials α_k and β_l and, on entering them into the corresponding cells, we get Table 9. We then find the indirect costs c'_{pq}; entering them in the corresponding cells we get Table 10.

Table 8

α \ β				
	1	2		
		3	2	
			2	1

Table 9

β α	0	1	0	−1
1	1	2		
2		3	2	
2			2	1

Table 10

β α	0	1	0	−1
1	1	2	1	$0'$
2	$2'$	3	2	$1'$
2	$2'$	$3'$	2	1

In fact, the three elementary stages we have considered could be united into one.

We now compute the differences $s_{pq} = c_{pq} - c'_{pq}$. In this case

$$\begin{aligned} s_{13} &= 3 - 1' = 2, \\ s_{14} &= 4 - 0' = 4, \\ s_{21} &= 4 - 2' = 2, \\ s_{24} &= 0 - 1' = -1, \\ s_{31} &= 0 - 2' = -2, \\ s_{32} &= 2 - 3' = -1 \end{aligned}$$

Hence the function S expressed in terms of nonbasic unknowns assumes the form

$$S = 420 + 2x_{13} + 4x_{14} + 2x_{21} - x_{24} - 2x_{31} - x_{32} \tag{8}$$

Among the coefficients of the unknowns in the right member there are negative ones, the coefficient x_{31}, for example. It is therefore possible to decrease the value of S by increasing x_{31} (while preserving the zero values of the remaining nonbasic unknowns).

We set $x_{31} = \rho$. Since the sums of the values of the unknowns in the rows and columns must remain unchanged, it is necessary to add ρ to or subtract it from the values of the basic unknowns in the following manner:

Table 11

$40-\rho$	$20+\rho$		
	$40-\rho$	$40+\rho$	
ρ		$40-\rho$	60

We had to compensate for the addition of ρ to x_{31} by subtracting ρ from x_{11} and in turn to compensate for this by the addition of ρ to x_{12} and so on until we came back to x_{31}.

We stop here to make a remark concerning Table 11. Going in the table from cell to cell in that sequence in which we compensate for ρ we get a closed broken line consisting of alternating horizontal and vertical links: the broken line is dotted in Table 11. One of the vertices of the broken line is in a nonbasic cell (x_{31} in this case) and the rest are in basic cells (not necessarily in all of them; thus there is no vertex in the cell x_{34} of Table 11). This broken line is called a *cycle* or, to be more exact, a *recalculation cycle* corresponding to a given nonbasic cell.

As is seen from Table 11, the condition that the unknowns should be nonnegative allows ρ to be increased only to 40. We set $\rho = 40$; three basic unknowns, x_{11}, x_{22} and x_{33}, then vanish together. We choose one of them, x_{22}, for example. So we introduce the unknown x_{31} into the basis and make the unknown x_{22} nonbasic. The new basic solution will be as follows:

$X_2 =$

0	60		
		80	
40		0	60

the value of the function S for it being

$$420 - 2 \times 40 = 340$$

(in (8) we set $x_{31} = \rho = 40$ and assume the values of the remaining unknowns to be zero). Going over to the new basis completes one step in the procedure. We have succeeded in decreasing the value of the function S by 80 units.

To sum up, by the beginning of the first step we had been given an initial basic solution $\underset{1}{X}$. The step itself consisted in going over to new basic solution $\underset{2}{X}$ and included the following stages:

(1) finding the potentials α_k, β_l and the indirect costs

$$c'_{pq} = \alpha_p + \beta_q$$

(2) computing the differences $s_{pq} = c_{pq} - c'_{pq}$;

(3) choosing a vacant cell corresponding to a negative difference s_{pq} and constructing the recalculation cycle for this cell;

(4) finding the new basic solution and the new value of the function S.

Notice that had all the differences s_{pq} proved nonnegative, that would have meant that the basic solution $\underset{1}{X}$ was optimal Then stages (3), (4) would have become unnecessary. Thus the *nonnegativity of all differences* s_{pq} *is the criterion for terminating the procedure.*

Let us return to the example under consideration, however. The next step begins with finding the potentials, indirect costs and computing the differences s_{pq}. The corresponding computation will look as follows:

β α	0	1	2	1
1	1	2	3′	2′
0	0′	1′	2	1′
0	0	1′	2	1

$$s_{13} = 3 - 3' = 0 \qquad s_{22} = 3 - 1' = 2$$
$$s_{14} = 4 - 2' = 2 \qquad s_{24} = 0 - 1' = -1$$
$$s_{21} = 4 - 0' = 4 \qquad s_{32} = 2 - 1' = 1$$

$$S = 340 + 2x_{14} + 4x_{21} + 2x_{22} - x_{24} + x_{32}$$

There are negative differences among s_{pq}: $s_{24} = -1$. We fix on the cell x_{24} and construct a recalculation cycle for it:

0	60		
		$80 - \rho$	ρ
40		$0 + \rho$	$60 - \rho$

We conclude from the examination of the cycle that ρ can be increased only to 60. Setting $\rho = 60$, we arrive at a new basic solution:

$\underset{3}{X} =$

0	60		
		20	60
40		60	

The value of the function S is

$$340 - 1 \times 60 = 280$$

for this solution. Going over from $\underset{2}{X}$ to $\underset{3}{X}$ completes the second step in the procedure.

The third step again begins with computing the potentials, indirect costs and differences s_{pq}. We have

β α	0	1	2	0
1	1	2	$3'$	$1'$
0	$0'$	$1'$	2	0
0	0	$1'$	2	$0'$

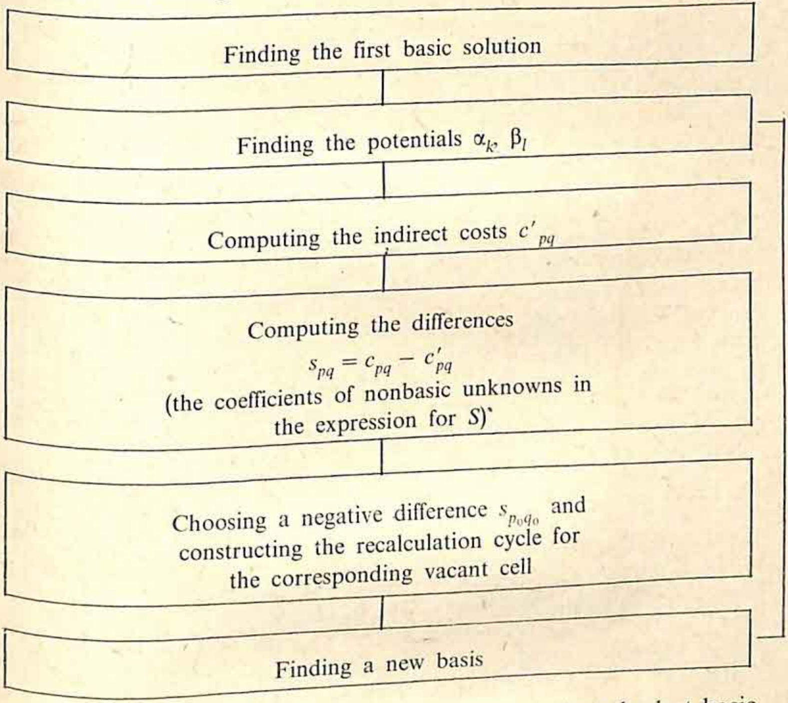
$$\begin{array}{ll} s_{13} = 3 - 3' = 0 & s_{22} = 3 - 1' = 2 \\ s_{14} = 4 - 1' = 3 & s_{32} = 2 - 1' = 1 \\ s_{21} = 4 - 0' = 4 & s_{34} = 1 - 0' = 1 \end{array}$$

Now the differences s_{pq} are all nonnegative. This means that the last basic solution $\underset{3}{X}$ is optimal and the corresponding (smallest) value of the function S is 280.

Thus the optimal solution of the problem has been found:

$$\begin{cases} x_{12} = x_{24} = x_{33} = 60, \quad x_{23} = 20, \quad x_{31} = 40; \\ \text{the remaining } x_{ij} \text{ are zero.} \end{cases}$$

In conclusion we show once again, but now only schematically, the plan of solving the transportation problem:


Finding the first basic solution
Finding the potentials α_k, β_l
Computing the indirect costs c'_{pq}
Computing the differences $s_{pq} = c_{pq} - c'_{pq}$ (the coefficients of nonbasic unknowns in the expression for S)
Choosing a negative difference $s_{p_0q_0}$ and constructing the recalculation cycle for the corresponding vacant cell
Finding a new basis

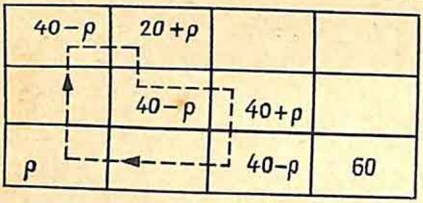
Note. If the differences s_{pq} are all nonnegative, then the last basic solution is optimal.

4°. *Justifying the method of potentials*. The formula

$$s_{pq} = c_{pq} - c'_{pq} \tag{7}$$

allowing the function S to be expressed in terms of nonbasic unknowns is the pivot of the solution method described above. Deriving this formula requires a deeper study of the transportation table and we have agreed to discuss it separately. We shall do it now. To that end we return to the study of Table 11

Table 11



$40-\rho$	$20+\rho$		
	$40-\rho$	$40+\rho$	
ρ		$40-\rho$	60

which contains the recalculation cycle for the cell x_{31}. We used this table to find ρ and thereby go over to a new basic solution. But the table allows still another problem to be solved: it helps to determine what are the coefficients of the unknown x_{31} in the expressions for the basic unknowns.

We first of all make the following remark. If the values of nonbasic unknowns are indicated, then this unambiguously determines the values of the basic unknowns. But it is just Table 11 that gives an idea of what happens to the values of basic unknowns when the nonbasic unknown x_{31} is assigned a value ρ (while assuming the remaining nonbasic unknowns to be zero). The study of the table shows that x_{31} enters the expressions for the basic unknowns having the following coefficients:

$$\left.\begin{array}{l} -1 \text{ in } x_{11} \\ +1 \text{ in } x_{12} \\ -1 \text{ in } x_{22} \\ +1 \text{ in } x_{23} \\ -1 \text{ in } x_{33} \end{array}\right\} \qquad (9)$$

In the remaining basic unknowns (namely, in x_{34}) x_{31} has the coefficient 0. We see that when making the round of the cycle the coefficients involved alternatively assume the values $+1$, -1.

It now remains to answer the question: What coefficient has x_{31} in the expression for the function S in terms of nonbasic unknowns? Well, this expression

$$S = 420 + s_{13}x_{13} + s_{14}x_{14} + s_{21}x_{21} + s_{24}x_{24} + \\ + s_{31}x_{31} + s_{32}x_{32}$$

is produced when the basic unknowns in the original expression for S

$$S = \Sigma c_{ij}x_{ij} \quad (i = 1, 2, 3; j = 1, 2, 3, 4)$$

are substituted for by their expressions in terms of nonbasic ones. In the given case it is necessary that $x_{11}, x_{12}, x_{22}, x_{23}, x_{33}, x_{34}$ should be substituted for by their expressions in terms of $x_{13}, x_{14}, x_{21}, x_{24}, x_{31}, x_{32}$. In virtue of (9) the total coefficient of x_{31} will be

$$s_{31} = c_{31} - c_{11} + c_{12} - c_{22} + c_{23} - c_{33}$$

Hence

$$s_{31} = c_{31} - (\alpha_1 + \beta_1) + (\alpha_1 + \beta_2) - (\alpha_2 + \beta_2) +$$
$$+ (\alpha_2 + \beta_3) - (\alpha_3 + \beta_3) = c_{31} - (\alpha_3 + \beta_1)$$

or

$$s_{31} = c_{31} - c'_{31}$$

But that is just formula (7) (for the given case). The reasoning that has led us to the last equality is sufficiently demonstrative, and there is no need to repeat it for the general case.

We conclude by saying a few words about other methods of solving the transportation problem. The method of potentials discussed above tends to be used in hand computation since one of its stages, namely that of constructing the recalculation cycle, is difficult to realize in the computer. With a small number of suppliers and consumers (within some ten) this method leads to an optimal solution in a reasonable time. If, however, the number of suppliers and that of consumers are large enough (and this is what happens in applications, where these numbers are of an order of hundreds), resorting to a computer is inevitable. Algorithms computer programmes are based on considerations other than those underlying the method of potentials. The interested reader can find more about these algorithms in special literature devoted to the transportation problem.

TO THE READER

Mir Publishers would be grateful for your comments on the content, translation and design of this book.

We would also be pleased to receive any other suggestions you may wish to make.

Our address:
USSR, 129820, Moscow,
I-110, GSP, Pervy Rizhsky
Pereulok, 2
Mir Publishers

Printed in the Union of Soviet Socialist Republics

METHOD OF SUCCESSIVE APPROXIMATIONS

N. VILENKIN, D.SC.

This book explains in a popular form the methods of approximation. solutions of algebraic, trigonometric, model and other equations.

Intended for senior schoolchildren, polytechnic students, mathematics teachers and for those who encounter solutions of equations in their practical work. In the course of exposition some elementary ideas about higher mathematics are introduced in the book. About 20 solved exercises are included in the appendix.

The book has already been translated into Spanish and French.

BASIC COMPUTATIONAL MATHEMATICS

V. DYACHENKO, D.SC.

The present text covers the most elementary concepts and ideas underlying modern computational methods for solving problems in mechanics and mathematical physics. It also deals with the construction and investigation of appropriate computational algorithms.

The presentation is simple and does not require that the reader have a substantial mathematical background. The text is designed for students of natural-science departments of colleges and universities. It will also be of interest to a broad range of physicists and engineers serving as an introductory course in computational mathematics.

FUNDAMENTAL THEOREM OF ARITHMETIC

L. KALUZHNIN, D.SC.

This booklet is devoted to one of the fundamental propositions of the arithmetic of rational whole numbers—the unique factorisation of whole numbers into prime multipliers. It gives a rigorous and complete proof of this basic fact. It is shown that uniqueness of factorisation also exists in arithmetic of complex (Gaussian) whole numbers. The link between arithmetic of Gaussian numbers and the problem of representing whole numbers as sum of squares is indicated. An example of arithmetic in which uniqueness of expansion into prime multipliers does not hold is given. The booklet is intended for the senior schoolchildren. It will help to acquaint them with the elements of number theory. It may also be useful for secondary school teachers.

INDUCTION IN GEOMETRY

L. GOLOVINA, D. SC. and I. YAGLOM, D. SC.

This booklet deals with various applications of the method of mathematical induction to solving geometric problems and was planned by the authors as a natural continuation of I. S. Sominsky's booklet "The Method of Mathematical Induction" published by Mir Publishers in 1975. It contains 37 worked examples and 40 problems accompanied by brief hints.

The book tells about the relation of systems of linear inequalities to convex polyhedra, gives a description of the set of all solutions of a system of linear inequalities, analyses the questions of compatibility and incompatibility; finally, it gives an insight into linear programming as one of the topics in the theory of systems of linear inequalities. The last section but one gives a proof of the duality theorem of linear programming.

The book is intended for senior pupils and all amateur mathematicians.

Mir Publishers
Moscow